



UNIVERSIDAD DE SANTIAGO DE COMPOSTELA
DEPARTAMENTO DE GENÉTICA

**DESARROLLO DE UN MAPA GENÉTICO CON
MARCADORES AFLP Y MICROSATÉLITE EN
RODABALLO (*Scophthalmus maximus* L.)**

Memoria presentada por:

Gloria M^a González Fortes

para optar al grado de Doctora en Ciencias Biológicas

Lugo, Mayo 2008

CARMEN BOUZA FERNÁNDEZ, PAULINO MARTÍNEZ PORTELA y LAURA SÁNCHEZ PIÑÓN, Profesores Titulares y Catedrática del Departamento de Genética de la Universidad de Santiago de Compostela

CERTIFICAN

Que el trabajo titulado “Desarrollo de un mapa genético con marcadores AFLP y microsatélite en rodaballo (*Scophthalmus maximus* L.)”, que presenta Gloria M^a González Fortes para optar al Grado de Doctora en Ciencias Biológicas, ha sido realizado bajo su dirección, que lo consideran concluído y autorizan su presentación al tribunal calificador.

Y para que así conste, firman la presente en Lugo, a 26 de Mayo de 2008.

Dra. Carmen Bouza Fernández Dr. Paulino Martínez Portela

Dra. Laura Sánchez Piñón

Gloria M^a González Fortes

Este trabajo ha sido financiado mediante el Proyecto de Investigación “Desarrollo de un mapa genético de ligamiento con marcadores microsatélite en rodaballo (*Scophthalmus maximus* L.)”, (PGIDT01MAR26101PR), concedido por la Xunta de Galicia. Así mismo, su realización se ha visto facilitada gracias al disfrute de una Bolsa de Tercer Ciclo y una Bolsa Predoctoral concedidas por la Consellería de Educación y Ordenación Universitaria (Dirección Xeral de Universidades e Investigación) de la Xunta de Galicia.

AGRADECIMIENTOS

Terminado el trabajo, queda mirar atrás, hacer un repaso e intentar no olvidar a ninguno de los que de una manera u otra me han ayudado a iniciar, desarrollar y concluir esta tesis.

En primer lugar, mi agradecimiento a mis directores de tesis, Carmen Bouza, Paulino Martínez y Laura Sanchez. Laura y Carmen confiaron en que podría llevar a cabo este proyecto y a ellas debo agradecerles el haberme iniciado en el mundo de la investigación. Especialmente, a Paulino y Carmen, les agradezco que hayan seguido este trabajo durante los últimos años, que han sido duros y a veces parecieron interminables.

A Francesco Nonnis-Marzano y Gilberto Gandolfi, de la Università degli Studi di Parma, debo agradecerles todo el trabajo relacionado con el desarrollo de marcadores AFLP, que fueron fundamentales para la elaboración de este mapa. Les debo mucho a nivel profesional, pero más aún a nivel personal. En su laboratorio comprobé que se puede disfrutar trabajando.

Gracias al Dr. Paolo Ajmone-Marsan della Università del Sacro Coure descubrí el GenomiPhi, que sirvió para rectificar los errores de principiante.

A Ania, Maria Peque y María Portela quiero agradecerles su compañía, los ratos de cháchara y de trabajo en equipo...de ellas me llevo los mejores recuerdos del laboratorio en Lugo. También a Belén y Sonia, les agradezco su profesionalidad y que en los momentos peores hayan tenido siempre una palabra de ánimo. Gracias también a Miguel por su ayuda con el misterioso mundo de los softwares estadísticos.

A mis compañeros en Italia tengo mucho que agradecerles. Marta Scalfi me mostró los secretos del Mapmaker, recordaré siempre la tarde que pasamos a los pies de los Alpes, portatil en mano, descifrando el lenguaje de Lander *et al.*, 1987. A Andrea y Pio, gracias por esos cafés que nos reunían a la comunidad de becarios para "fregarsene di tutto". A Milena, Riccardo y Marina, siempre les agradeceré el haberme recibido como una más en su laboratorio, sin mostrar un mínimo de recelo ante la española que monopolizaba sus pipetas.

A mi familia le agradezco su apoyo, el que siempre estén ahí y me valoren igual con tesis que sin ella. A Marco, que comenzó conmigo la aventura lucense y que ha vivido de cerca los alti-bajos de estos años, gracias por tu confianza y apoyo incondicional.

No quiero cerrar estos agradecimientos sin mencionar a mis amigas Ángeles, Amalia, Eda ..., gracias a ellas los largos inviernos lucenses fueron menos fríos.

RESUMEN

RESUMEN

A lo largo de esta memoria se presenta la construcción de un mapa genético en una de las especies marinas de mayor interés en la acuicultura gallega, el rodaballo (*Scophthalmus maximus*). En especies de interés comercial, la disponibilidad de un mapa genético supone una herramienta fundamental en la mejora de los stocks de cultivo, ya que son el primer paso en la identificación de QTLs (*Quantitative Trait Loci*) para caracteres de interés como la resistencia a enfermedades, crecimiento o determinación sexual. El desarrollo de estos mapas genéticos exige disponer de familias adecuadas para el seguimiento de la segregación génica y de un número de marcadores relativamente elevado que permitan cubrir el genoma en estudio. En el caso de rodaballo, una especie de domesticación relativamente reciente y poco conocida a nivel genómico, cuando se emprendió este proyecto de mapeo, no se disponía de familias seleccionadas y el número de marcadores desarrollados era escaso. Así, la selección de la familia de referencia para el mapeo y la identificación de marcadores de ADN que cubrieran adecuadamente el genoma, marcaron gran parte del desarrollo de este trabajo de Tesis.

En rodaballo se disponía de la tecnología para la inducción de ginogénesis, por lo que se seleccionó como familia de mapeo una hembra de rodaballo procedente de una población natural y 50 de sus embriones ginogenéticos haploides, considerando las ventajas de los individuos haploides en cuanto al genotipado de marcadores dominantes. En una muestra de estos embriones haploides se puso a punto, por primera vez en esta especie, la tecnología AFLP. A partir de la combinación de endonucleasas *EcoRI*/*TaqI* y de diecisiete parejas de cebadores, se identificaron 186 fragmentos polimórficos en la familia de mapeo. Setenta y cuatro polimorfismos fueron descartados atendiendo a criterios de calidad de la amplificación y de ajuste a la ratio de segregación mendeliana 1:1 ($P < 0,01$), y finalmente, 112 marcadores

AFLP se utilizaron en el análisis de ligamiento. Entre las bandas con ratios distorsionadas y las descartadas por su escasa calidad, se estudió el efecto de posibles amplificaciones inespecíficas y la presencia de bandas complejas en la inducción de errores de genotipado.

Además de los marcadores AFLP se incluyeron en la construcción del mapa 24 marcadores microsatélite: 22 procedentes de la literatura y 2 puestos a punto en este trabajo a partir de una genoteca de rodaballo enriquecida en secuencias microsatélite. El mapa resultante incluye 111 marcadores (91 AFLPs y 20 microsatélites) distribuidos en 23 grupos de ligamiento (LG) -uno más que el número de cromosomas de la especie- que abarcan un total de 1040 cM, siendo la distancia media entre marcadores adyacentes de 14,6 cM. Los marcadores se distribuyeron aleatoriamente en el mapa, sin que pudieran identificarse agrupaciones significativas. A partir de los marcadores *framework* se obtuvo una primera estima del tamaño genómico total de rodaballo, que sería de 1150 cM.

A pesar de las limitaciones de un mapa de moderada densidad como el presentado en este estudio, sus grupos de ligamiento proporcionan un primer marco de organización del genoma de rodaballo, dibujando los primeros bocetos de los cromosomas físicos subyacentes al mapa genético. La amplificación de los marcadores AFLP y microsatélite de este estudio en nuevas familias segregantes junto con nuevos marcadores permitirá el anclaje a grupos de ligamiento de mapas genéticos futuros hasta consolidar un mapa genómico de referencia en esta especie.

SUMMARY

This work reports the construction of the genetic map for one of the most important marine species in the Galician aquaculture, the turbot (*Scophthalmus maximus*). In species with commercial value, the availability of a genetic maps is an important tool for breeding, as it is the first step for QTL (*Quantitative Trait Loci*) identification of traits of productive interest, such as disease resistant loci. The construction of genetics maps needs appropriate families to follow the genetic marker segregation, and a relative high number of markers to cover the genome under study. In the case of turbot, a species of recent domestication and with a low knowledge about its genome, at the beginning of the present mapping project, there was a lack of selected families for linkage analysis and the number of available markers was low. Thus, the selection of the reference family for mapping and the identification of DNA markers for an appropriate coverage of the turbot genome were major underlying specific objectives to the development of this PhD work.

Using the available technology for gynogenesis induction in this species, a turbot female from a wild population and its progeny of 50 gynogenetic embryos were selected as family mapping, taking into account the advantages of the haploid individuals for dominant markers genotyping. In a sample of these embryos, the AFLP technology was applied for the first time in turbot. From the combination of endonucleases *EcoRI/TaqI* and seventeen primer pairs, 186 polymorphic fragments were identified in the mapping family. Seventy-four were discarded attending to quality criterion and the distorted *ratios* from mendelian segregation 1:1 ($P < 0.01$). Finally, 112 AFLP markers were used for the linkage analysis. The effect of mismatch and size homoplasia in genotyping errors were studied between the rejected fragments.

In addition to the AFLP markers, 24 microsatellites were included in the map construction: 22 from the literature and two from a microsatellite enriched

genomic library developed in this work. The resulting map includes 111 markers (91 AFLPs and 20 microsatellites) distributed over 23 linkage groups (LGs), one more than expected attending to the haploid number of turbot. The map spans 1040 cM, with a mean distance between adjacent markers of 14.6 cM. Markers appeared randomly distributed over the genome. A total map length about 1150 cM in turbot was estimated from the framework markers in this mapping project.

This moderately dense genetic map represents the first step in the knowledge of the turbot genome organization, drawing the outline of the physical chromosomes under the genetic map's LGs. The characterization of the AFLP and microsatellite markers from this study in other mapping families with new markers will be able to anchor the present linkage groups with those of future maps to consolidate a reference genomic map in this species.

ÍNDICE

I. INTRODUCCIÓN.....	1
I.1. PLEURONECTIFORMES.....	3
I.1.1. Sistemática y principales características morfológicas y biológicas de los peces planos.....	3
I.1.2. El rodaballo (<i>Scophthalmus maximus</i>).....	5
I.1.2.1. Sistemática.....	5
I.1.2.2. Características biológicas.....	6
I.1.2.3. Producción de rodaballo.....	7
I.1.2.4. Problemática asociada al cultivo de rodaballo.....	9
I.1.2.5. Mejora genética en rodaballo.....	11
I.2. MAPAS GENÉTICOS.....	13
I.2.1. Principios básicos en el mapeo genético.....	14
I.2.1.1. Las leyes de la herencia.....	14
I.2.1.2. Ligamiento incompleto y recombinación.....	14
I.2.1.3. Entrecruzamiento, recombinación y distancia de mapa....	16
I.2.1.4. Entrecruzamiento múltiple, análisis de tres-puntos e interferencia.....	18
I.2.1.5. Funciones de mapa.....	19
I.3. CONSTRUCCIÓN MAPAS GENÉTICOS.....	22
I.3.1. Poblaciones de referencia en la construcción de mapas de ligamiento.....	23
I.3.1.1. Mapeo genético en cruces experimentales.....	23
I.3.1.2. Mapeo genético en poblaciones no manipuladas.....	24
I.3.1.3. Líneas ginogenéticas.....	26
I.3.2. Marcadores moleculares para el mapeo genómico.....	27
I.3.2.1. Características generales de los marcadores genéticos.....	27
I.3.2.2. Marcadores específicos de secuencia. Microsatélites.....	32
I.3.2.3. Marcadores independientes de secuencia. AFLP.....	35

I.3.3. Análisis computacional del ligamiento genético.....	38
I.3.3.1. Análisis de los loci individuales.....	38
I.3.3.2. Análisis de dos puntos.....	39
I.3.3.3. Análisis multipunto.....	40
I.4. MAPEO GENÓMICO EN PECES.....	41
I.5. OBJETIVOS.....	44
II. MATERIAL Y MÉTODOS.....	47
II. 1. MATERIAL BIOLÓGICO.....	49
II.1.1. Inducción de ginogénesis.....	49
II.1.2. Extracción de ADN a partir de familias ginogenéticas haploides.....	50
II.1.2.1. Extracción de ADN genómico de alto peso molecular en rodaballos diploides adultos.....	51
II.1.2.2. Extracción de ADN genómico de alto peso molecular en embriones haploides de rodaballo.....	51
II.1.3. Selección de la familia de referencia.....	52
II.1.4. Amplificación genómica global. Genomi Phi™.....	54
II.2. DESARROLLO Y CARACTERIZACIÓN DE MARCADORES MOLECULARES.....	56
II.2.1. Marcadores microsatélite.....	56
II.2.1.1 Estrategias para la obtención de marcadores microsatélite.....	58
II.2.1.1.1. Amplificación cruzada.....	58
II.2.1.1.2 Librerías enriquecidas.....	60
II.2.1.2. Puesta a punto de las condiciones de amplificación PCR	61

II.2.1.3. Identificación de loci polimórficos.....	61
II.2.1.3.1. Detección automatizada de fragmentos de ADN....	61
II.2.1.4 Evaluación de la diversidad genética.....	63
II.2.1.5. Genotipado de marcadores microsatélite en la familia de referencia.....	63
II.2.2. Tecnología AFLP.....	64
II.2.2.1. Optimización de la metodología AFLP.....	64
II.2.2.2 Protocolo AFLP.....	65
II.2.2.3 Selección de cebadores AFLP.....	66
II.2.2.4 Detección automatizada de AFLPs.....	67
II.2.2.5. Análisis de los perfiles AFLP. Software GENOGRAPHER	68
II.2.2.6 Nomenclatura AFLP.....	70
II.2.3. Test de conformidad con la segregación mendeliana.....	71
II.2.4. Fuentes de error de genotipado.....	71
II.2.4.1. Precauciones metodológicas en la caracterización de marcadores AFLP.....	73
II.2.4.2. Dos fuentes particulares de errores de genotipado: amplificaciones inespecíficas y presencia de bandas complejas.	74
II.2.4.2.1. Evaluación de posibles amplificaciones inespecíficas.....	74
II.2.4.2.2. Caracterización de bandas complejas. Homoplasia	75
II.3. CONSTRUCCIÓN DEL MAPA GENÉTICO DE LIGAMIENTO.....	76
II.3.1. Softwares de análisis de ligamiento.....	76
II.3.2. MAPMAKER/EXP 3.0.....	78
II.3.2.1. Codificación para MAPMAKER de los datos genotípicos procedentes de familias ginogenéticas haploides.	80
II.3.3 Identificación de los grupos de ligamiento. Análisis de dos puntos.....	82

II.3.3.1 Estimación de la frecuencia de recombinación. Método de Máxima Verosimilitud.....	82
II.3.3.2 Determinación de ligamiento entre pares de marcadores: Test de razón de verosimilitud y método de puntuación LOD...	84
II.3.3.3. Criterios de asignación de ligamiento.....	85
II.3.3.4. Tamaño muestral y precisión de la estima de r	86
II.3.4 Ordenación de loci: análisis multipunto.....	88
II.3.4.1 Análisis de tres puntos.....	88
II.3.4.2 Estimación de la tasa de recombinación multilocus.....	89
II.3.4.3 Ordenación de los loci dentro de sus grupos de ligamiento.....	89
II.3.5 Detección de errores de genotipado: algoritmo de Lincoln y Lander (1992).....	92
 II.4. ESTUDIO DE LA DISTRIBUCIÓN DE MARCADORES A LO LARGO DEL GENOMA.....	94
II.4.1. Distribución de marcadores dentro de los grupos de ligamiento.....	94
II.4.2. Distribución de marcadores a lo largo del mapa.....	94
II.4.2.1. Evaluación de la distribución de marcadores independientemente de su composición.	95
II.4.2.2. Evaluación de la distribución de marcadores en función del tipo de marcador y de su composición.....	95
 II. 5. ESTIMA DE LA LONGITUD GENÓMICA Y COBERTURA DEL MAPA DE <i>Scophthalmus maximus</i>	96
 III. RESULTADOS.....	99
III.1. MATERIAL GINOGENÉTICO HAPLOIDE.....	101
III.1.1. Verificación de ginogénesis.....	101

III.1.2. Familia de referencia.....	102
III.1.3. Obtención de ADN a partir de embriones haploides. Rendimiento del protocolo de extracción estándar vr. protocolo de extracción optimizado.....	103
III.1.4. Amplificación genómica.....	104
III. 2. MARCADORES MOLECULARES: Microsatélites y AFLPs.....	105
III.2.1. Marcadores microsatélite.....	105
III.2.1.1. Nuevos loci microsatélite en <i>Scophthalmus maximus</i>	105
III.2.1.1.1. Identificación de secuencias microsatélite. Eficiencia de las distintas estrategias de búsqueda.....	105
III.2.1.1.2. Diversidad genética de los nuevos microsatélites	107
III.2.1.2. Variabilidad de los marcadores microsatélite en la familia de referencia.....	108
III.2.1.3. Relación de ligamiento genético entre los loci microsatélite.....	108
III.2.2 Marcadores AFLP.....	110
III.2.2.1. ADN genómico.....	110
III.2.2.2. Perfiles AFLP.....	111
III.2.2.3. Productividad de los cebadores AFLP.....	112
III.2.2.4. Nivel de polimorfismo detectado.....	114
III.2.2.5. Relación entre composición nucleotídica y productividad de los cebadores.....	115
III.2.3. Caracterización de las meiosis informativas. Análisis de la distorsión de segregación.....	116
III.2.4. Identificación de fuentes de error de genotipado en marcadores AFLP.....	117
III.2.4.1. Identificación de amplificaciones inespecíficas.....	118
III.2.4.2. Identificación de bandas complejas.....	120

III.3. CARTOGRAFÍA GENÉTICA DEL RODABALLO.....	126
III.3.1. Análisis dos puntos. Asignación de los marcadores a los grupos de ligamiento.....	124
III.3.2 Análisis multipunto. Orden de marcadores y distancias de mapa.....	125
III.3.3. Test de asociación entre LGs e identificación de genotipos erróneos.....	127
III.3.4. Características del mapa de ligamiento genético de <i>S. maximus</i>	127
III.3.5. Presencia en el mapa de marcadores con distorsión de segregación.....	129
III.4. DISTRIBUCIÓN DE LOS MARCADORES A LO LARGO DEL MAPA..	130
III.4.1. Distribución de marcadores dentro de los grupos de ligamiento.....	130
III.4.2 Distribución de marcadores entre los grupos de ligamiento.....	131
III.4.2.1. Distribución de marcadores independientemente de su origen y composición.....	131
III.4.2.2. Distribución de marcadores en función de su origen y composición nucleotídica.....	132
III.5. LONGITUD GENÓMICA Y COBERTURA DEL MAPA FEMENINO DE <i>S. maximus</i>	133
IV. DISCUSIÓN.....	137

IV.1. ESTRATEGIA DE MAPEO: FAMILIAS DE GINOGENÉTICOS HAPLOIDES.....	139
IV.2. HERRAMIENTAS MOLECULARES: MARCADORES MICROSATÉLITE Y AFLP.....	142
IV.2.1. Marcadores microsatélite en <i>Scophthalmus maximus</i>	144
IV.2.2. Marcadores AFLP en <i>Scophthalmus maximus</i>	147
IV.2.3. Algunas consideraciones acerca de los marcadores microsatélite y AFLP en el análisis de ligamiento.....	157
IV. 3. MAPA GENÉTICO DE <i>S. maximus</i>	159
IV.3.1. Ligamiento genético entre pares de marcadores.....	159
IV.3.2. Ordenación de los marcadores en los grupos de ligamiento.....	161
IV.3.3. Distancia de mapa.....	164
IV.3.4. Distribución de los marcadores.....	165
IV.3.5. Distorsión de segregación.....	170
IV. 3.6. Longitud genómica y cobertura del mapa.....	171
IV. 3.7. Distancia física vs. distancia de mapa.....	173
V. CONCLUSIONES.....	179
VI. BIBLIOGRAFÍA.....	185
VII. APÉNDICE.....	213
VII. ABREVIATURAS.....	225

I. INTRODUCCIÓN

I.1. PLEURONECTIFORMES.

I.1.1. Sistemática y principales características morfológicas y biológicas de los peces planos.

Los peces planos, incluido el rodaballo (*Scophthalmus maximus*), pertenecen al grupo de los teleósteos y se clasifican dentro del orden Pleuronectiformes.

Etimológicamente, la denominación de este orden procede de las palabras griegas *pleura* (lado) y *nekton* (nadador) y del término latino *forma*. Así, el nombre hace referencia a una de las características más notables de este grupo de peces: los individuos adultos presentan un cuerpo lateralmente comprimido y asimétrico, adaptado a la vida sobre el fondo bentónico.

La asimetría corporal de los Pleuronectiformes se debe a una metamorfosis que tiene lugar durante los primeros estadios de desarrollo larvario. Originalmente, las larvas son simétricas y nadan con normalidad, sin embargo, en el transcurso de su desarrollo uno de los ojos comienza a migrar hacia el lado opuesto, desplazándose sobre la cabeza o perforando los tejidos de la parte superior hasta colocarse al lado del ojo que ha permanecido fijo. La migración ocular va acompañada de una torsión de la parte superior del cuerpo que da lugar a la asimetría. Desde ese momento, los peces jóvenes nadan sobre un flanco. Dependiendo del flanco en que se posicionen los ojos, las especies pleuronectiformes se definen como orientadas a la derecha o a la izquierda (Figura I.1).

La singular conformación corporal de este grupo de peces está relacionada con sus costumbres bentónicas. Los individuos adultos carecen de vejiga natatoria y viven en fondos blandos, compuestos de arena y limo, sobre los que se desplazan dejando ver sólo los ojos, en ocasiones protuberantes, como adaptación a la vida sobre el fondo marino.

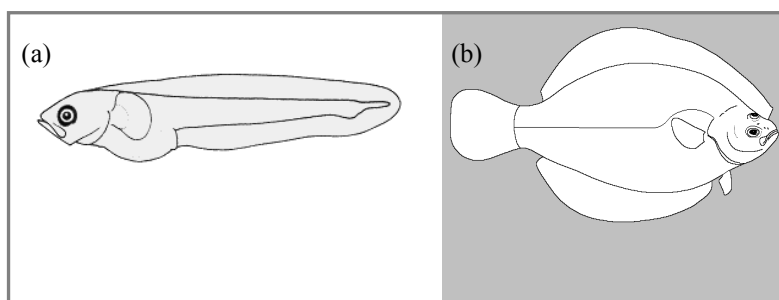


Figura I.1. (a) Larva y (b) adulto del orden Pleuronectiformes (FishBase; www.fishbase.org)

El flanco superior, sobre el que se sitúan los ojos, generalmente presenta manchas, diseños punteados, jaspeados o veteados y esta lleno de cromatóforos, lo que permite a estos peces adaptarse a la coloración del sustrato en respuesta a estímulos visuales, variando su color entre el gris, pardo, bruno y verdusco. Por el contrario, el flanco ciego presenta un aspecto blanquecino, desprovisto de pigmentación.

Otra de las adaptaciones a la vida bentónica se refleja en las aletas caudales y dorsales, que se expanden ampliamente sobre los costados dorsal y ventral hasta cubrir casi por completo el perímetro del animal.

Al orden Pleuronectiformes pertenecen más de 600 especies distribuidas en catorce familias (Nelson, 1994). De ellas, las más conocidas son las que habitan en las costas marinas y que tradicionalmente se han explotado para la pesca comercial. Sin embargo, existen también Pleuronectiformes abisales y tropicales e incluso algunas especies de agua dulce, difundidas principalmente en Asia y América meridional. Es en las aguas tropicales de la región Indo-Pacífica donde se ha descrito una mayor diversidad dentro del grupo de peces planos, sin embargo, el pequeño tamaño de estos peces y el desconocimiento de muchos de los hábitats tropicales dificultan su caracterización, por lo que estas especies siguen siendo muy poco conocidas, tanto a nivel taxonómico como ecológico (Gibson, 2005).

Entre los pleuronectiformes marinos que habitan las aguas del hemisferio norte, se encuentran algunas de las especies más valiosas por su importancia económica para la pesca comercial y la acuicultura. Además de la solla (*Pleuronectes platessa*) y el rodaballo (*Scophthalmus maximus*), se incluyen en este grupo el fletán (*Hippoglossus hippoglossus*) y la limanda (*Limanda limanda*) de los mares del norte.

I.1.2. El rodaballo (*Scophthalmus maximus*)

I.1.2.1. Sistemática

Dentro de los Pleuronectiformes, el rodaballo pertenece a la familia Scophthalmidae incluida en el suborden de los Pleuronectoidei (Gibson, 2005; Pardo *et al.*, 2005). En cuanto al género, existe cierta controversia a la hora de clasificar al rodaballo dentro del género *Scophthalmus* (Wheeler, 1992; Desoutter *et al.*, 2001; Chanet, 2003) o *Psetta* (Nielsen, 1986), ambos pertenecientes a la familia Scophthalmidae. Actualmente se admiten ambas clasificaciones taxonómicas (Fishbase; <http://www.fishbase.org/>), dependiendo el uso de las denominaciones *Scophthalmus maximus* (Linnaeus, 1758) o *Psetta maxima* (Linnaeus, 1758) de las preferencias de cada autor.

En el presente estudio se ha optado por la nomenclatura *Scophthalmus maximus*, siguiendo la propuesta de Wheeler (1992). Datos de distancia genética basados en isoenzimas (Blanquer *et al.*, 1992; Bouza *et al.*, 1997), así como la alta similitud a nivel cariotípico demostrada entre rodaballo y otras especies pertenecientes al género *Scophthalmus* (Bouza *et al.*, 1994a; Pardo *et al.*, 2001) apoyan esta clasificación. Igualmente, la hibridación natural entre las dos especies del género *Scophthalmus*, rémol (*S. rhombus*) y rodaballo (*S. maximus*) (Purdom, 1976), apunta en la misma dirección. Consecuentemente, la taxonomía del rodaballo sería la siguiente:

Supertaxa: Pisces
Superclase: Osteichthyes
Clase: Actinopterygii
Subclase: Neopterygii
Infraclasse: Teleostei
Superorden: Acanthopterygii
Orden: Pleuronectiformes
Suborden: Pleuronectoidei
Familia: Scopthalmidae
Género: Scopthalmus
Especie: *Scopthalmus*
maximus

I.1.2.2. Características biológicas

El rodaballo es una especie marina bentónica que vive en fondos arenosos, rocosos o mixtos, en un rango de profundidad de 20 a 100 m, siendo los individuos más jóvenes los que suelen habitar las áreas menos profundas.

Presenta una amplia distribución geográfica a lo largo de las costas del Atlántico nororiental: desde Marruecos recorre toda la costa europea hasta el Círculo Polar Ártico y el Mar Báltico; en el área mediterránea se encuentra tanto en las costas norafricanas como en las del sur europeo (Blanquer *et al.*, 1992). En el mar Negro se ha descrito la existencia de un taxón morfológicamente diferenciado, considerado por algunos autores como subespecie (*Psetta maxima maeotica* ; Jones, 1974). Sin embargo, estudios alozímicos poblacionales han evidenciado una escasa diferenciación genética del taxón *maeotica* respecto a otras poblaciones de rodaballo del Mediterráneo occidental (Blanquer *et al.*, 1992).

S. maximus comparte las características morfológicas de los Pleuronectiformes, con un cuerpo casi circular, plano y asimétrico (Figura I.2). Los ojos son protuberantes y se sitúan sobre el flanco izquierdo. La piel no tiene escamas, si bien la cara dorsal presenta unas protuberancias óseas irregularmente distribuidas.

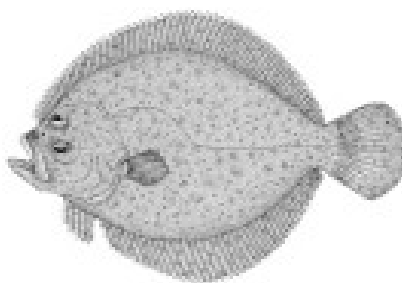


Figura I.2. Rodaballo adulto (Cultured Aquatic Species Information Programme, FAO)

Son animales carnívoros, los ejemplares jóvenes se alimentan de bivalvos y crustáceos, mientras que el alimento principal de los individuos adultos son otros peces del fondo (lanzones, gobios, etc.) y cefalópodos.

No presentan dimorfismo sexual, pero los machos son de menor tamaño y crecen más lentamente. Alcanzan la madurez sexual entre el tercer y quinto año de vida, dependiendo de la temperatura del agua. En la etapa adulta, algunos ejemplares pueden alcanzar tallas máximas de 25 kg y 1 m de longitud.

El desove ocurre entre Febrero y Abril en el área mediterránea y entre Mayo y Agosto en el Atlántico. Concretamente en Galicia, la freza tiene lugar en los meses de Mayo, Junio y Julio, respondiendo a un fotoperíodo creciente y a un aumento de la temperatura del agua. Los huevos son pelágicos y presentan una única gota de grasa. Tras un período de incubación de 5 a 7 días eclosionan las larvas, de vida pelágica y simetría bilateral. La migración ocular que dará lugar a la fase post-larvaria tiene lugar aproximadamente a los 50 días de la eclosión.

I.1.2.3. Producción de rodaballo.

El rodaballo es un pez de alto valor económico de carne blanca y blanda muy apreciada en alimentación. Su producción en cautividad es una práctica relativamente reciente. Los primeros cultivos de esta especie se llevaron a cabo en el Reino Unido a comienzos de los años 70 (Purdom *et al.*,

1972) y no se introdujeron en otros países europeos hasta comienzos de los 80. Estos primeros cultivos se centraban en el engorde de alevines obtenidos a partir de poblaciones naturales. Posteriormente, se consiguió evitar la dependencia del medio natural al optimizar la cría larvaria en las “*hatcheries*” o instalaciones de cría y cerrar así el ciclo de cultivo en cautividad (Devesa, 1995).

Tradicionalmente la producción de rodaballo se ha concentrado en los países de la costa Atlántica europea, donde se ha consolidado una industria moderna que actualmente lidera el mercado mundial de esta especie (Figura I.3). Otros países productores fuera del área de distribución de la especie son Chile, donde el rodaballo se introdujo con éxito en los años 80 y más recientemente China, donde el cultivo de esta especie ha experimentado un rápido desarrollo (Jilin, 2003).

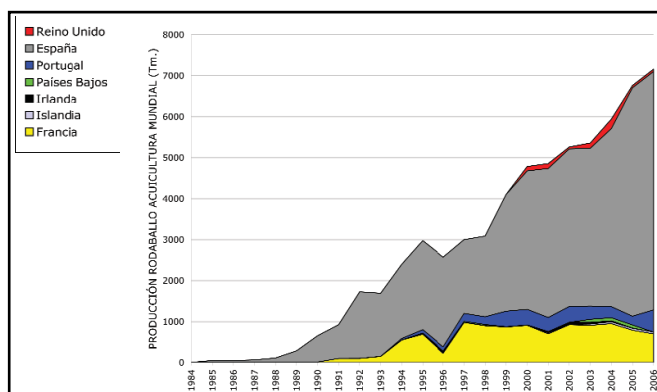


Figura I.3. Evolución de la producción acuícola de rodaballo en Europa para el periodo 1980-2006 (FAO+FEAP).

El cultivo de rodaballo comenzó en España en los años 80 y actualmente nuestro país lidera la producción mundial de rodaballo de crianza, aportando un 81,2% del rodaballo de piscifactoría de toda Europa (FEAP, 2006). La producción se lleva a cabo en Galicia, Cantabria y País Vasco, pero es la comunidad gallega la que produce la práctica totalidad de los alevines para las granjas y aporta el 92% de la producción nacional de rodaballo. La calidad y la temperatura del agua en las costas gallegas han permitido liderar el sector, al resultar especialmente adecuadas para el engorde de esta especie

(Slasky, 1985; Bouza *et al.*, 1997). En las granjas gallegas, los ejemplares de rodaballo llegan a alcanzar la talla comercial en 2 o 3 años, mientras que en costas más frías como las de Francia, Noruega o Reino Unido son necesarios 3 y 4 años para alcanzar el mismo peso.

En 2006, el rodaballo de piscifactoría supuso el 52,5% del total de rodaballo comercializado en Europa. El resto fue aportado por la pesca extractiva –procedente principalmente de los Países Bajos-, que en esta especie sigue abasteciendo una parte importante del mercado. Sin embargo, el sector de la pesca ha entrado en declive, lo que contrasta con el crecimiento continuo de la producción industrial (Figura I.4).

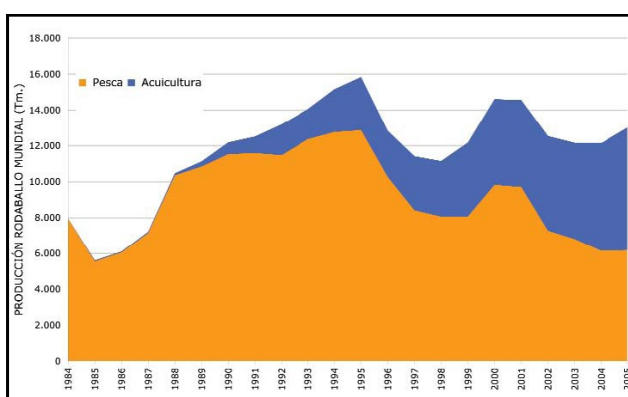


Figura I.4. Evolución de las fuentes de obtención de rodaballo en Europa: Acuicultura vs. Pesca extractiva para el período 1980-2005 (FAO).

Ante esta perspectiva, los futuros incrementos en la demanda comercial de rodaballo deberán cubrirse con la producción procedente de la acuicultura. Este mercado en expansión plantea al sector de la acuicultura una serie de retos que deberá superar para alcanzar un volumen de producción que asegure su competitividad.

I.1.2.4. Problemática asociada al cultivo de rodaballo.

El aumento de la producción de rodaballo supone solventar algunas de las dificultades con las que se ha venido enfrentando desde sus comienzos: aumentar la viabilidad larvaria; acortar el tiempo de desarrollo necesario para que los ejemplares alcancen la talla comercial; controlar la incidencia de

enfermedades asociadas al cultivo intensivo del rodaballo; y evitar los problemas derivados de la depresión endogámica.

La baja viabilidad larvaria es una de los problemas con los que tradicionalmente se encontrado el cultivo de rodaballo. Tan sólo un 10-20% de los individuos superan la fase larvaria (Devesa, 1995), siendo el inicio de la alimentación tras la reabsorción del saco vitelino (2-7 días) y la transición de rotífero a artemia (12-15 días) las fases más críticas para la supervivencia (Rodríguez y Díaz, 1990). Actualmente, se ha conseguido reducir la mortalidad ejerciendo un mayor control sobre la calidad del agua, los niveles de oxígeno en los tanques, la alimentación y parámetros físico-químicos como luz y temperatura, aunque la viabilidad sigue estando por debajo de la observada en otras especies cultivadas de peces planos.

Superada la metamorfosis, el tiempo necesario para que los juveniles de rodaballo alcancen el tamaño comercial es de 26 a 30 meses. En ese período apenas hay problemas de mortalidad, pero el cultivo se enfrenta a una importante dispersión de tallas. La diversidad de tamaño de los individuos obliga a clasificarlos en diferentes tanques en función de su talla, para evitar que se establezcan jerarquías y ajustar el tamaño del alimento a la boca de los peces. El factor más influyente en este crecimiento diferencial es un marcado dimorfismo sexual en la tasa de engorde. Las hembras de rodaballo tienen una tasa de crecimiento que es más de dos veces la del macho (Husebye *et al.*, 1994), que crece más lentamente y alcanza antes la madurez sexual, lo que disminuye su crecimiento somático, al tiempo que aumenta su mortalidad y el riesgo de padecer enfermedades (Imsland *et al.*, 1997; Piferrer *et al.*, 2000). Las características más favorables de las hembras en el período de engorde han llevado a un creciente interés en la obtención de poblaciones femeninas para el cultivo (Piferrer *et al.*, 2003). Sin embargo, se desconoce todavía cuál es el sistema de determinación sexual en esta especie.

En cuanto a las enfermedades, la producción de rodaballo se ha visto en ocasiones mermada debido a septicemias causadas por *Vibrio* y *Pseudomonas*, o a otras enfermedades de origen fúngico, vírico o parasitario. Además de medidas de higiene relacionadas con el control de la calidad del agua en los tanques, en los últimos años se han llevado a cabo programas de vacunación en los alevines, que han conseguido disminuir la mortalidad ligada a enfermedades (Devesa, 1995) y mejorar la calidad del producto final.

1.1.2.5. Mejora genética en rodaballo

En general, la problemática asociada a la acuicultura y en particular al cultivo del rodaballo, se ha abordado fundamentalmente desde el estudio de la fisiología de los peces cultivados y la optimización de las instalaciones, adecuándolas a las necesidades de cada especie. Sólo en los últimos años, una vez establecidos los parámetros biológicos y técnicos del cultivo, la industria piscícola ha comenzado a invertir esfuerzos en desarrollar programas de mejora genética que puedan dar respuesta a la problemática asociada con la producción.

Estos programas de mejora combinaban los principios de la genética cuantitativa con la capacidad de asignación de parentesco de los marcadores moleculares. Los primeros esfuerzos se centraron por tanto, en la caracterización de marcadores en rodaballo, que posteriormente se utilizarían para evaluar la diversidad genética existente en las poblaciones naturales y de piscifactoría (Bouza *et al.*, 1994b; 1997; 2002), así como para el desarrollo y seguimiento de programas de selección de reproductores partiendo de stocks genéticamente diferenciados (Castro *et al.*, 2004).

Además de las prácticas tradicionales de mejora genética, en rodaballo se han aplicado de forma experimental técnicas de manipulación cromosómica, obteniendo con éxito individuos triploides y ginogenéticos haploides y diploides (Piferrer *et al.*, 2000; 2003; 2004). La inducción de la triploidía es una práctica habitual en cultivos marinos, ya que las poblaciones resultantes están

constituidas por individuos estériles que no sufren los problemas asociados con la maduración sexual (Thorgaard, 1983; 1986; Piferrer *et al.*, 1995; Donaldson *et al.*, 1996; Arai, 1999; Benfey, 1999). En el caso de la inducción de ginogénesis, el principal objetivo es la obtención de poblaciones femeninas, caracterizadas por elevadas tasas de crecimiento (Piferrer *et al.*, 2003).

En los últimos años, el pujante interés en el sector acuícola ha favorecido una importante inversión económica para el desarrollo de nuevas herramientas relacionadas con la mejora genética, entre ellas destaca la elaboración de mapas genéticos en las principales especies de la acuicultura: salmón atlántico (*Salmo salar*, Gharbi *et al.*, 2006), lubina (*Dicentrarchus labrax*, Chistiakov *et al.*, 2005), dorada (*Sparus aurata*, Franch *et al.*, 2006) y la carpa común (*Cyprinus Carpio*, Sun y Liang, 2004). Estos mapas han permitido identificar relaciones de ligamiento entre marcadores conocidos y loci asociados con caracteres cuantitativos (QTL: *Quantitative Trait Loci*) de interés económico, como la resistencia a enfermedades o la tolerancia a determinadas temperaturas (Ozaki *et al.*, 2001; Cnaani *et al.*, 2003), lo que los ha convertido en valiosas herramientas para la selección asistida por marcadores (Danzmann y Gharbi, 2001; Yu y Guo, 2006). Además de estas aplicaciones, los mapas genéticos ofrecen una importante información acerca de la organización y evolución del genoma en un grupo tan heterogéneo como el de los peces, al tiempo que posibilitan el mapeo comparativo con otras especies de vertebrados (Naruse *et al.*, 2000; Jaillon *et al.*, 2004; Kai *et al.*, 2005).

En el caso de rodaballo, disponer de un mapa genético de moderada densidad podría ser útil en la selección asistida por marcadores a favor de determinados QTLs de interés comercial, como por ejemplo, el crecimiento, la determinación sexual, la resistencia a patologías o anomalías de la pigmentación o (Raimuzzo *et al.*, 1994; Wang *et al.*, 2006; Cnaani *et al.*, 2003; 2007). Además, un mapa genético de esta especie ayudaría a caracterizar la organización de su genoma, hasta ahora prácticamente desconocida. La

información disponible acerca del genoma de rodaballo procede fundamentalmente de estudios citogenéticos, que han establecido que el cariotipo de rodaballo esta constituido por 22 pares de cromosomas (20 acrocéntricos y 2 metacéntricos), sin mostrar heteromorfismos cromosómicos asociados al sexo (Bouza *et al.*, 1994a; Pardo *et al.*, 2001; Cuñado *et al.*, 2001; 2002). En este sentido, el mapa genético podría contribuir a definir el sistema de determinación sexual en rodaballo, en la medida que puedan identificarse en esta especie marcadores asociados al sexo.

I.2. MAPAS GENÉTICOS

El mapa genético o mapa de ligamiento de una especie es una representación del orden lineal de un grupo de genes y/o marcadores a lo largo del genoma, siendo la distancia entre loci adyacentes proporcional a la frecuencia de recombinación entre los mismos (Danzmann y Gharbi, 2001). Los marcadores ordenados se distribuyen en grupos de ligamiento, cuyo número debe corresponder al número haploide de cromosomas de la especie cuando el mapa tiene una cobertura genómica adecuada.

En un proyecto genómico, estos mapas de ligamiento facilitan el marco de referencia para el ensamblaje a lo largo de los cromosomas de las secuencias brutas de ADN obtenidas en los programas de secuenciación genómica. Sin embargo, los mapas genéticos proporcionan por sí mismos importante información acerca de los genomas: son esenciales para entender el comportamiento de los cromosomas y sus interacciones durante la meiosis (Sakamoto *et al.*, 2000); permiten la localización de elementos genómicos estructurales, marcadores moleculares y loci que influyen en caracteres mendelianos pero también la disección genética de rasgos cuantitativos (QTL) (Lynch y Walsh, 1998); suministran estimas del tamaño genómico; y posibilitan el estudio evolutivo mediante mapeo comparativo. Por ello, en la mayoría de las especies, incluyendo los peces, uno de los principales objetivos de los

proyectos genómicos ha sido la obtención de mapas de ligamiento (Kocher *et al.*, 1998; Young *et al.*, 1998; Shimoda *et al.*, 1999; Naruse *et al.*, 2000).

I.2.1. Principios básicos en el mapeo genético

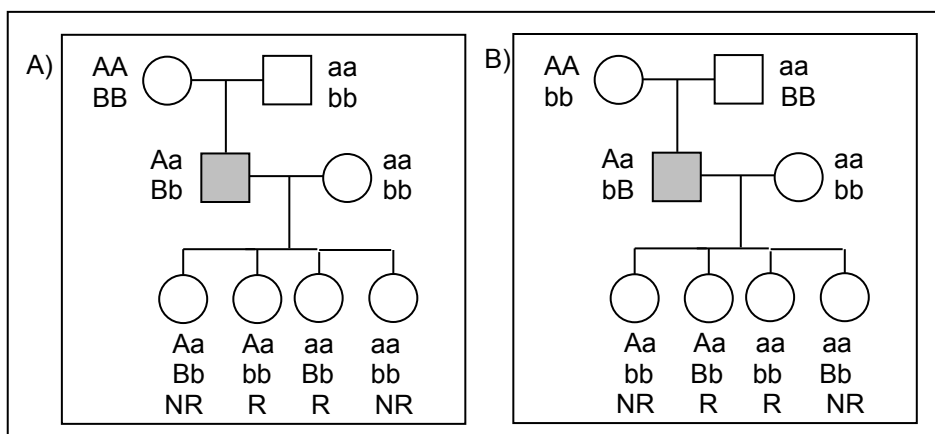
I.2.1.1. Las leyes de la herencia

“Durante la formación de los gametos, la segregación de un par de genes es independiente de la de otros pares de genes”. Este principio se conoce generalmente como tercera ley de Mendel o principio de la transmisión independiente de los caracteres.

A principios del siglo XX, W. Bateson, R.C. Punnet y T.H. Morgan detectaron desviaciones de la tercera ley de Mendel. Estas excepciones son la base biológica para el mapeo genético. Independientemente uno de otro, Bateson y Punnet en el guisante dulce y Morgan en *Drosophila melanogaster*, encontraron que en cruzamientos dihíbridos algunos genotipos aparecían con más frecuencia que otros, indicando que los pares de genes no se transmitían de manera independiente. Los haplotipos parentales de la generación P (parental), pasaban preferentemente juntos a la descendencia de la generación F₁ (filial). Este fenómeno se definió como ligamiento. Para explicar esta segregación no independiente, Morgan asumió que los genes que se heredan juntos están localizados en el mismo cromosoma, por lo que tienden a transmitirse unidos más frecuentemente que si estuvieran en cromosomas diferentes. Los estudios de Morgan junto con la observación del comportamiento de los cromosomas durante la meiosis proporcionaron las bases para la teoría cromosómica de la herencia, según la cual los genes se disponen de forma lineal en los cromosomas.

I.2.1.2. Ligamiento incompleto y recombinación

De acuerdo con la teoría cromosómica de la herencia, los genes situados en diferentes cromosomas segregarán de forma independiente y aquellos en el mismo cromosoma mostrarán ligamiento.



15

Considerando el ligamiento incompleto resultante del entrecruzamiento, Morgan estableció que los genes con frecuencias de recombinación menor al 50% se encuentran en el mismo cromosoma (ligados). Mientras que los genes con segregación independiente mostrarán una frecuencia de recombinación igual al 50%.

1.2.1.3 Entrecruzamiento, recombinación y distancia de mapa

La ocurrencia del entrecruzamiento a lo largo de un par de cromosomas es “semialeatoria” (Watson *et al*, 1987), siendo este principio la base del mapeo génico. Cuanto más próximos estén dos loci, menor es la probabilidad de que ocurra un entrecruzamiento entre ellos. A medida que aumenta la distancia entre dos loci, la oportunidad de sufrir un entrecruzamiento aumentará. Estas observaciones permitieron a Morgan usar el número de entrecruzamientos, estimado a través de la frecuencia de recombinación entre dos loci (r), como una medida de la distancia entre los mismos. Así, la distancia de mapa entre dos loci se conoce como unidad de mapa o centimorgan (cM), de tal forma que 1 unidad de mapa es igual a 1 cM y corresponde a un 1% de recombinación entre los genes (Figura. 1.6).

Alfred H. Sturtevant (1913), discípulo de Morgan, fue el primero que usó la variación en la intensidad del ligamiento para determinar la secuencia de genes en la dimensión lineal del cromosoma, desarrollando así el concepto de mapa genético. Sturtevant sostuvo que si el orden de los genes en el cromosoma es lineal y las frecuencias de recombinación dependen de las distancias físicas entre ellos, entonces los genes podrían ordenarse como puntos en una línea recta en distancias proporcionales a la frecuencia de recombinación (Figura 1.7). Un mapa genético se puede extender siguiendo este razonamiento hasta incluir todos los genes conocidos en un cromosoma; estos genes constituirán un grupo de ligamiento. El número de grupos de ligamiento en un mapa debe ser el mismo que el número haploide de cromosomas de la especie.

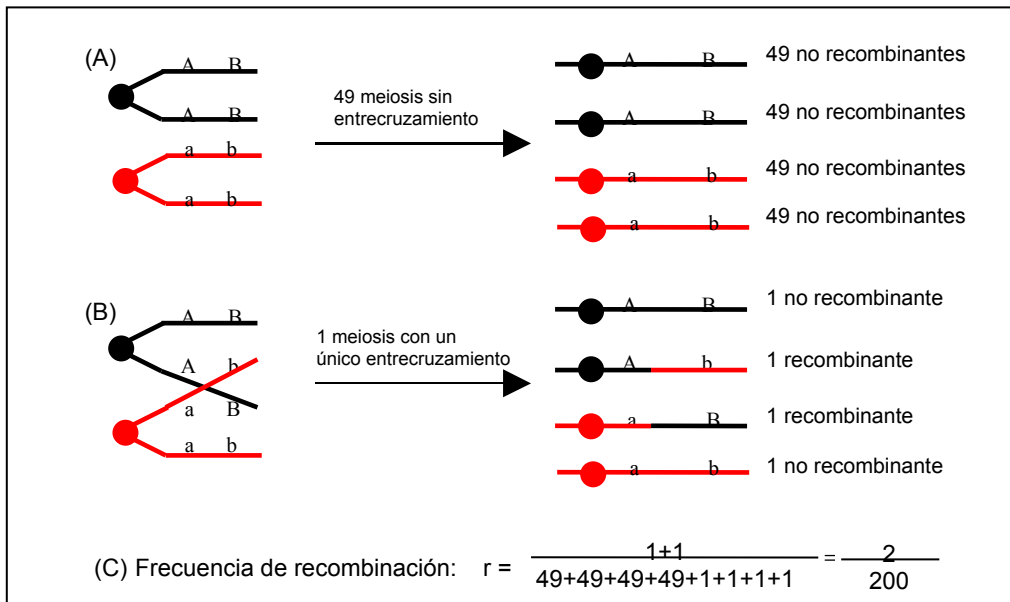


Figura. I.6. Físicamente, 1 unidad de mapa corresponde a la longitud del cromosoma en la cual, como media, tiene lugar un entrecruzamiento por cada 50 células que sufren meiosis. Si hay un entrecruzamiento en una célula meiótica de 50, la frecuencia de entrecruzamiento es $1/50$, o 2%. Sin embargo, la frecuencia de recombinación entre los genes será del 1%. Una frecuencia de entrecruzamiento del 2% significa que de 200 cromosomas que resultan de la meiosis en 50 células, exactamente 2 cromosomas son recombinantes para los marcadores genéticos que abarcan el particular segmento cromosómico. Un 2% de entrecruzamiento corresponde a un 1% de recombinación, puesto que sólo la mitad de las cromátidas en cada célula con un intercambio son de hecho recombinantes. Modificado de Hartl y Jones (1998).

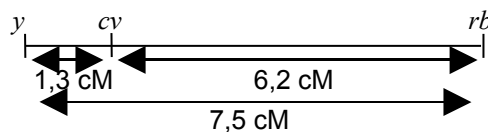


Figura I.7. Mapa genético para genes situados en el cromosoma X de *Drosophila*: y (cuerpo amarillo), rb (ojos rubí) y cv (alas miniatura o *crossvein*). Dadas las frecuencias de recombinación observadas $r_{y,rb} = 7.5\%$, $r_{rb,cv} = 6.2\%$ y $r_{y,cv} = 1.3\%$ el mapa genético correcto será aquel que respete la aditividad de las frecuencias de recombinación entre los genes ($r_{y,cv} \approx r_{y,rb} + r_{rb,cv}$)

1.2.1.4. Entrecruzamiento múltiple, análisis de tres-puntos e interferencia

Cuando la distancia física entre dos loci es lo suficientemente grande, pueden ocurrir dos (o incluso más) entrecruzamientos en la región que los separa. Si los entrecruzamientos implican al mismo par de cromátidas no hermanas, el resultado es que hay un intercambio físico de un segmento cromosómico entre los loci, pero no se detecta recombinación entre los pares de alelos, puesto que el segundo entrecruzamiento revierte el efecto del primero (Figura 1.8). Dado que los dobles entrecruzamientos en principio no son detectables en las progenies, el valor de recombinación observado entre pares de loci será una infraestima de la frecuencia de intercambio real y por tanto, de la distancia de mapa entre los loci.

Para detectar los dobles entrecruzamientos es preciso seguir la segregación de un tercer locus en el estudio de la recombinación (Figura 1.8). Estos cruces de tres puntos incluyen un par alélico localizado en el espacio interior de los marcadores genéticos, de tal forma que los intercambios dobles en la región podrán ser detectados cuando los entrecruzamientos flanqueen a este gen. La identificación de los tipos doble recombinantes convierte a los cruces de tres puntos en un método no sólo eficiente para obtener datos de recombinación, sino también para determinar el orden de los genes de un modo simple.

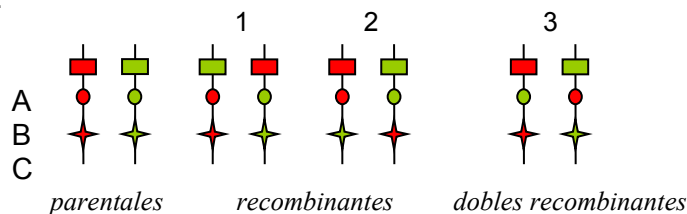


Figura. 1.8. Para los loci A, B y C ordenados en un par de cromosomas homólogos, se ilustran las disposiciones alélicas parentales y las resultantes de recombinações simples (1 y 2) y dobles (3). Cuando ocurre un único entrecruzamiento entre loci adyacentes (recombinantes 1 y 2), la recombinación es observable entre los loci extremos A y C. Sin embargo, cuando ocurre un doble entrecruzamiento (recombinante 3) la recombinación pasa inadvertida en un análisis de dos puntos externos, ya que la disposición de los alelos A y C es la parental

El análisis de estos cruces de tres puntos reveló además, que las recombinaciones simultáneas no son independientes, sino que, en general, las recombinaciones dobles tienen lugar con una frecuencia menor de la esperada bajo la asunción de independencia (Muller, 1916). Así, dados los marcadores ordenados A-B-C: $r_{AC} < r_{AB} + r_{BC}$. Esto es debido al efecto del fenómeno conocido como *interferencia cromosómica* (i), según el cual un entrecruzamiento en una región de un cromosoma reduce la probabilidad de un segundo entrecruzamiento en una región cercana. La magnitud de la interferencia a lo largo de un cromosoma se mide a partir del *coeficiente de coincidencia* (c), que es igual al cociente entre el número observado y el esperado de dobles recombinantes.

$$i = 1 - c$$

La interferencia influye en la relación entre distancia de mapa y frecuencia de recombinación y su valor varía dependiendo de la distancia entre marcadores y también de la especie, del sexo del individuo o de la región cromosómica. Generalmente, la interferencia aumenta a medida que disminuye la distancia entre marcadores, reduciendo la probabilidad de los dobles entrecruzamientos, hasta minimizar el coeficiente de coincidencia que se igualaría a cero. En esta situación (generalmente por debajo de las 10 unidades de mapa (cM)) las frecuencias de recombinación entre los genes son aditivas y los mapas podrían extenderse sumando las distancias entre loci adyacentes hasta incluir todos los loci conocidos de un cromosoma. Para distancias mayores a 10 cM, la interferencia disminuye y el coeficiente de coincidencia va aumentando hasta un máximo de uno. En esta situación, la frecuencia de recombinación observada es menor que la distancia de mapa. La magnitud de tal diferencia dependerá del patrón de interferencia a lo largo del cromosoma.

I.2.1.5. Funciones de mapa

La variación en el patrón de interferencia a lo largo de los cromosomas complica la relación inicialmente intuitiva entre la probabilidad de que se

verifique un evento de recombinación y la distancia entre marcadores. Esta complicación ha llevado al desarrollo de expresiones matemáticas conocidas como funciones de mapa, que traducen la frecuencia de recombinación observada entre dos loci en la correspondiente distancia de mapa en cM. Cada patrón de interferencia rinde una función de mapa diferente.

Cuando se puede excluir la ocurrencia de múltiples entrecruzamientos entre dos loci, es decir, cuando la interferencia es completa ($i = 1$) la función de mapa es simple: la distancia de mapa (x) es igual a la fracción de recombinación (r):

$$x = r$$

Esta suposición está garantizada para genes próximamente ligados (generalmente, cuando $r < 10\%$). Cuando en un cromosoma se dispone de muchos loci estrechamente ligados y se conoce su orden, el método más simple para calcular las distancias de mapa entre ellos es estimar la fracción de recombinación en cada intervalo de loci adyacentes. La distancia de mapa entre dos loci distantes se obtiene entonces como la suma de las distancias de mapa de los intervalos entre estos loci (Sturtevant, 1913). La ecuación se conoce también como función de mapa de Morgan (Morgan, 1928).

Cuando entre los loci ocurren múltiples entrecruzamientos, las fracciones de recombinación en los diferentes intervalos dejan de ser aditivas. Bajo la asunción de que los entrecruzamientos en diferentes intervalos ocurren de acuerdo a la distribución de Poisson, Haldane (1919) expresó la distancia de mapa (x) como

$$x = -1/2 \ln(1-2r) \quad (1.1)$$

cuya inversa es

$$r = 1/2 (1 - e^{-2x}) \quad (1.2)$$

La función de mapa de Haldane ignora el fenómeno de la interferencia y asume que $i = 0$.

Dependiendo de la asunción del mecanismo por el cual opera la interferencia, se han derivado muchas otras funciones de mapa. Particularmente, en la función de mapa de Kosambi (1944) se asume que la interferencia disminuye como una función lineal de la distancia de acuerdo a $i = 1 - 2r$ y se deriva la función de mapa como:

$$\chi = \frac{1}{4} \ln (1+2r/1-2r) \quad (1.3)$$

cuya inversa es

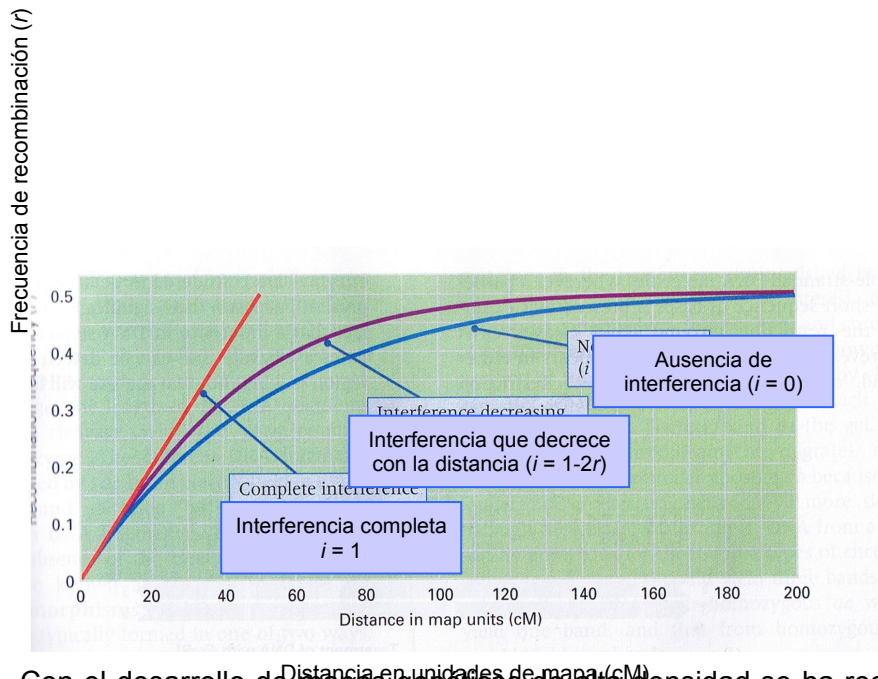
$$r = \frac{1}{2} (e^{4\chi} - 1)/(e^{4\chi} + 1) \quad (1.4)$$

Debido a la asunción de interferencia, los mapas de ligamiento contruidos usando la función de Kosambi son más cortos que los contruidos usando la función de Haldane.

La Tabla I.1 incluye algunas de las funciones de mapa más usadas. Cada función asume diferentes patrones de interferencia, desde 0 en la función de Haldane hasta interferencia completa en la función de Morgan. En una representación lineal, cualquiera de estas funciones discurrirá en el intervalo entre ambas relaciones extremas (Figura I.9)

Tabla I.1. Funciones de mapa ordenadas por el nivel de interferencia asumido.

Función de mapa	Coefficiente de coincidencia	Distancia de mapa como función de r
Haldane (1919)	1	$-1/2 \ln(1-2r)$
Kosambi (1944)	$2r$	$1/4 \ln(1+2r)/(1-2r)$
Carter y Falconer (1951)	$8r^3$	$1/4 (\tan^{-1}2r + \tanh^{-1}2r)$
Morgan (1928)	0	r



Con el desarrollo de mapas genéticos de alta densidad se ha reabierto el debate científico acerca de la idoneidad de las funciones de mapa. La mayoría de las funciones desarrolladas calculan la distancia entre genes o marcadores considerando una diferencia que depende de la interferencia que se genera a lo largo del cromosoma. (Modificada de Hartl y Jones, 2002)

en cuenta otros factores que pueden afectar a la frecuencia de recombinación, como la localización de estos loci en el cromosoma, el sexo del individuo investigado o la especie a la que pertenece. Así, las funciones de mapa parecen funcionar sólo en condiciones determinadas de número de loci y de distancia entre los mismos, sin haber una función de mapa universal (Zhao y Speed, 1996; Windemuth *et al.*, 1998). Estas consideraciones han de tenerse en cuenta a la hora de interpretar el tamaño físico real de los cromosomas a partir de distancias genéticas procedentes de un mapa de ligamiento basado en frecuencias de recombinación.

I.3. CONSTRUCCIÓN DE MAPAS GENÉTICOS

Para que un mapa genético sea útil en la identificación de asociación entre marcadores y genes de interés a partir de un número limitado de meiosis,

los marcadores deben distribuirse a lo largo del genoma en intervalos no mayores de 20 cM (Strachan y Read, 1999). De hecho, para la clonación posicional, los mapas de referencia han de ser mucho más densos, de hasta 1 cM de separación entre marcadores (Gyapay *et al.*, 1994; Strachan y Read, 1999). El trazado de estos mapas genéticos detallados supone: la selección o el desarrollo de poblaciones informativas para el mapeo; la caracterización de un elevado número de marcadores; y el posterior análisis de ligamiento a partir de los datos genotípicos.

I.3.1. Poblaciones de referencia en la construcción de mapas de ligamiento.

La población usada para el mapeo de genes se denomina comúnmente población de mapeo. La posibilidad de obtener poblaciones de mapeo a partir de cruces experimentales dependerá de las características de la especie estudiada y esta posibilidad determinará la estrategia de mapeo (Grattapaglia y Sederoff, 1994).

I.3.1.1. Mapeo genético en cruces experimentales.

En muchas especies de plantas y animales pueden llevarse a cabo cruces específicos para simplificar la detección de ligamiento y la estimación de su intensidad (Weir, 1996). Estos cruces se diseñan buscando maximizar en el parental o parentales la frecuencia de loci informativos, es decir, loci heterocigóticos que permitan identificar la procedencia de los alelos en la descendencia. En tales cruces, se conoce la fase de ligamiento *cis* o *trans* entre los genes y/o marcadores segregantes, lo que permite identificar con facilidad los tipos parentales y recombinantes. Además, los pedigrís suelen estar completos y prácticamente no faltan datos a excepción de posibles problemas técnicos que puedan surgir en el genotipado de algún individuo. Esta es una importante ventaja en comparación con los a veces incompletos pedigrís naturales.

Entre los diferentes tipos de cruces experimentales uno de los más informativos para el mapeo genético es el retrocruzamiento (*BC: backcross*), donde un parental heterocigótico procedente de la unión de dos líneas puras se cruza con un individuo homocigótico de la generación anterior (Danzmann y Gharbi, 2001). En la progenie derivada del cruce, los genotipos heterocigóticos u homocigóticos para cada marcador segregarán en una proporción 1:1, siendo fácilmente identificables los eventos de recombinación que hayan tenido lugar en el parental heterocigótico.

Otra de las estrategias más frecuentes es el uso de poblaciones F_2 derivadas de dos líneas consanguíneas, donde los marcadores codominantes segregan en proporciones 1:2:1 o 1:1:1:1 y los dominantes en proporción 3:1. A pesar de ser más complejas de analizar, estas poblaciones proporcionan casi el doble de información que los retrocruzamientos, puesto que los marcadores son segregantes en ambos parentales (Lander *et al.*, 1987).

En plantas diploides autofertilizadas, pueden usarse los individuos F_1 de un cruce entre parentales no relacionados para fundar una serie de líneas consanguíneas recombinantes (*Recombinant Inbred Lines*, RIL) por sucesivas generaciones de auto-polinización. Después de varias generaciones, los genotipos presentan una proporción muy elevada de homocigotos y cualquier RIL puede considerarse para el diseño de cruces experimentales.

1.3.1.2. Mapeo genético en poblaciones no manipuladas

En poblaciones naturales en las cuales no se pueden planificar los emparejamientos y no se dispone de líneas consanguíneas, la detección de ligamiento se complica debido al número de alelos segregantes por locus y parental, y al desconocimiento de la fase de ligamiento entre los loci (Maliepaard *et al.*, 1997). Con esta problemática se enfrenta el mapeo en humanos o en especies con ciclos reproductivos largos o en un estadio temprano de domesticación. En estos casos, la falta de cruces experimentales

altamente informativos generalmente se compensa con el estudio de varios núcleos familiares de dos o tres generaciones (Lander *et al.*, 1987).

El desarrollo de la tecnología PCR permitió establecer una estrategia de mapeo diferente al posibilitar el reconocimiento de alelos de distintos marcadores genéticos en células espermáticas (revisión en Arnheim *et al.*, 1990). El análisis de estas células haploides permite identificar los marcadores sin interferencia de la dominancia ni necesidad de cruzamientos prueba. Además, este método tiene el potencial de estudiar decenas e incluso cientos de meiosis en un único individuo, incrementando la precisión en la determinación de la frecuencia de recombinación y superando las limitaciones de tamaño y tiempo de los pedigrís clásicos. Desafortunadamente, esta estrategia sólo puede usarse para mapear marcadores de enfermedades si la mutación que causa la enfermedad ha sido caracterizada a nivel genético (Strachan y Read, 1999).

En la misma línea, en coníferas, algunos investigadores propusieron la construcción de mapas de un único individuo a partir del análisis de los megagametofitos haploides, que representan en estas plantas los productos de la segregación meiótica femenina (Tulseriam *et al.*, 1992). Posteriormente, esta estrategia se ha utilizado con éxito en varios proyectos de mapeo en vegetales (Plomion *et al.*, 1995; Paglia *et al.*, 1998; Gosselin *et al.*, 2002), si bien presenta la misma limitación que el mapeo de células espermáticas en cuanto a la imposibilidad de mapear caracteres fenotípicos.

En especies animales, el uso de ginogenéticos haploides o diploides y de líneas familiares androgenéticas ha resultado muy útil para el mapeo genético cuando no se dispone de líneas consanguíneas (Danzmann y Gharbi, 2001). Como los megagametofitos en plantas, las familias ginogenéticas o androgenéticas en animales, permiten seguir la segregación de los marcadores en los gametos procedentes de la hembra o del macho, respectivamente, e identificar los posibles entrecruzamientos en el parental informativo.

1.3.1.3. Líneas ginogenéticas

En 1990, Lewin *et al.*, inspirados en la reciente aplicación de la PCR en células espermáticas humanas (Li *et al.*, 1988), propusieron una estrategia de análisis de ligamiento basada en la PCR en oocitos no fertilizados que permitía el análisis de gametos haploides.

En algunos organismos, como los peces, los huevos pueden ser pseudofertilizados con espermia irradiado o de una especie heteróloga dando lugar a individuos ginogenéticos haploides, que representan una única meiosis materna sin contribución genética paterna (Lie *et al.*, 1994). Aunque los embriones haploides generalmente no son viables, continúan el desarrollo por varios días, incrementando el número de células con cada división mitótica. Así, la estrategia ginogenética ofrece algunas ventajas respecto al tipado del espermia: pueden realizarse tests de repetición en el mismo individuo y los marcadores que pueden mapearse no se restringen al ADN. Varios proyectos de mapeo en especies de peces han seguido la estrategia propuesta por Lie *et al.* (1994), por ejemplo, Postlethwait *et al.* (1994) en el pez cebra, Kocher *et al.* (1998) en tilapia o más recientemente Sun y Liang (2004) en la carpa común.

Para evitar los problemas de inviabilidad de los embriones haploides, puede inducirse la diploidía mediante la retención del segundo corpúsculo polar o inhibiendo la primera división mitótica a través de un choque térmico o de presión (Thorgaard, 1983; Felip *et al.*, 2001). Estos individuos ginogenéticos diploides son especialmente útiles para estimar distancias gen-centrómero, en un procedimiento de mapeo similar al aplicado en *Neurospora* (Barratt *et al.*, 1954; Perkins, 1962). En los individuos diploides resultantes de la fusión del segundo corpúsculo polar, los pares de cromosomas proceden de cromátidas hermanas. Bajo esta circunstancia, la proporción de descendencia heterocigótica puede usarse para estimar las distancias de mapa entre un marcador y su centrómero. Esta suposición es válida en la medida en que los marcadores estén lo bastante próximos al centrómero para no formar

entrecruzamientos múltiples. La localización de marcadores más distantes se lleva a cabo entonces mapeando dichos marcadores en relación a otros más próximos al centrómero.

1.3.2. Marcadores moleculares para el mapeo genómico

Las señalizaciones en un mapa geográfico pueden no tener valor en sí mismas, pero sin ellas el viajero estaría desorientado y difícilmente llegaría a su destino. Siguiendo este símil, en un mapa genético los marcadores, ya sean mutaciones génicas reconocibles a nivel fenotípico o cambios silenciosos de la secuencia de ADN, van a servir como puntos de referencia, generalmente de escasa importancia en cuanto a su propio significado biológico, pero fundamentales para llegar establecer la localización de un gen de interés. La frecuencia de recombinación entre los marcadores y dicho gen servirán para estimar las distancias de mapa e identificar el orden de los loci ligados en un mismo cromosoma. Así, los marcadores son la herramienta básica de cualquier proyecto de mapeo génico y de su abundancia genómica y calidad de su genotipado dependerá la fiabilidad del mapa genético obtenido.

Un marcador genético ha de ajustarse a la siguiente definición: debe ser un polimorfismo genético heredable, con una interpretación inequívoca en cuanto a su segregación y una alta reproducibilidad (Liu, 1998). Tradicionalmente, en los estudios de genómica se han usado tres tipos de marcadores génicos: marcadores morfológicos, marcadores proteicos y marcadores de ADN.

1.3.2.1. Características generales de los marcadores genéticos.

Los primeros estudios de mapeo se basaron en caracteres morfológicos discretos tales como la forma, el color o el tamaño (por ejemplo, el primer mapa de ligamiento de *Drosophila* elaborado por Sturtevant en 1913 o el del tomate por Butler en 1952). Tales fenotipos son útiles como marcadores génicos cuando el carácter morfológico identifica un único gen funcional. Sin

embargo, muchos caracteres morfológicos no tienen una base genética simple, sino que resultan de la interacción de múltiples genes, pudiendo además estar influidos por el ambiente, lo cual dificulta la interpretación de la variación fenotípica en términos genéticos (Falconer y Mackay, 1996; Martínez, 2005). Por otro lado, puesto que en una población sólo segregan unos pocos caracteres morfológicos, se necesitarían muchas poblaciones de mapeo para obtener un número razonable de polimorfismos (Liu, 1998).

A partir de los años 60, el análisis de caracteres bioquímicos, tales como el grupo sanguíneo y las isoenzimas, más fácilmente interpretables y menos complejos que los caracteres morfológicos, incrementó el abanico de marcadores genéticos. Sin embargo, el número de polimorfismos era todavía insuficiente para cubrir la demanda de marcadores generada por cualquier intento de mapeo del genoma completo de una especie. Incluso en los organismos más estudiados, los mapas construidos a partir de loci de efectos fenotípicos conocidos incluían largos espacios intergénicos, que no podrían ser mapeados por análisis de ligamiento, puesto que no había marcadores para cubrir esas regiones.

La demanda de marcadores por parte de los proyectos de mapeo encontró respuesta en la genética molecular, con el desarrollo de los primeros marcadores basados en la detección de polimorfismos a nivel del ADN (Danzmann y Gharbi, 2001; Liu y Cordes, 2004; Chistiakov *et al.*, 2005). Típicamente, un marcador de ADN es una pequeña región del genoma que muestra, en diferentes individuos dentro de una especie, un polimorfismo de secuencia no asociado con ninguna alteración fenotípica medible (Griffiths *et al.*, 2002). Estos marcadores tienen una serie de características favorables comparados con los caracteres morfológicos o proteicos: permiten revelar un mayor número de polimorfismos (puesto que se identifican también las mutaciones silenciosas y sinónimas escondidas en las proteínas); no están influenciados por factores ambientales o por interacciones génicas complejas;

y son considerados en general caracteres neutrales, con muy poco o ningún valor adaptativo (Ajmone-Marsan *et al.*, 2001). En el campo de la genómica, la accesibilidad a los polimorfismos del ADN supuso disponer, por primera vez, de un conjunto de marcadores suficientemente numerosos y espaciados a través de todo el genoma, aplicables a cualquiera de las especies estudiadas.

Los primeros en utilizar las diferencias en las secuencias de ADN como marcadores para el mapeo génico fueron Botstein y colaboradores en 1980. A través de la técnica de hibridación previamente descrita por Southern (1975), evidenciaron la existencia de polimorfismos en determinadas secuencias diana de los enzimas de restricción. Tales polimorfismos serían consecuencia de mutaciones silenciosas que pueden ocurrir con bastante frecuencia en el ADN, creando o destruyendo sitios particulares de corte. La alteración de las dianas enzimáticas originará diferencias en la longitud de los fragmentos de restricción, que se conocen como RFLPs (*Restriction Fragment Length Polymorphism*). En una población, la existencia de distintos patrones de corte enzimático resultarán en genotipos alternativos de presencia o ausencia para un fragmento de un determinado tamaño. Puesto que los RFLP siguen las leyes mendelianas de la herencia, los loci heterocigóticos identificados se usaron para construir con éxito el primer mapa de ligamiento genético en humanos basado en marcadores de ADN (Botstein *et al.*, 1980). Sin embargo, los RFLP presentaban dos limitaciones importantes: i) la tecnología original era bastante laboriosa y costosa en tiempo, dinero y requerimientos de ADN; y ii) la información que aportan es limitada, puesto que se trata de marcadores bialélicos.

La puesta a punto de la reacción en cadena de la polimerasa (PCR) a mediados de los años 80 revolucionó la genética molecular y particularmente la velocidad de desarrollo de nuevos marcadores de ADN (Figura I.10). La nueva tecnología proporcionaba un método alternativo al laborioso *Southern blot* para identificar variaciones nucleotídicas en pequeñas regiones del genoma.

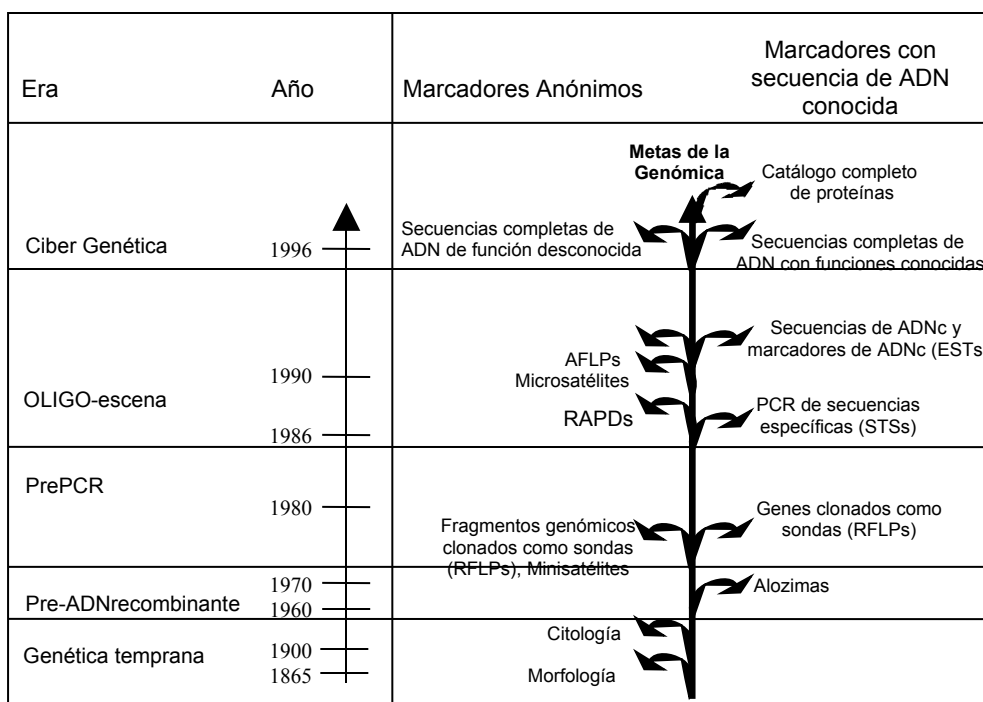


Figura I.10. “Árbol evolutivo” de los marcadores genéticos basado en un dibujo original de Dr. David O’Malley del Forest Biotechnology Group, North Carolina State University. Según Liu *et al.* (1998).

Los primeros marcadores basados en la PCR se desarrollaron a finales de los 80 y principios de los noventa: los RAPDs (*Random Amplified Polymorphic DNA*) y los microsatélites o SSR (*Simple Sequence Repeat*). La detección de los primeros se basa en la amplificación al azar de polimorfismos de ADN a partir de cebadores cortos de secuencia aleatoria. Los microsatélites, sin embargo, se basan en la amplificación de loci específicos a partir de cebadores complementarios a las secuencias que flanquean el segmento de interés. El número de marcadores basados en la PCR siguió aumentando a lo largo de la última década con la aparición de los AFLP (*Amplified Fragment Length Polymorphism*) y SNP (*Single Nucleotide Polymorphism*) entre otros. En pocos años, estos marcadores han ido desplazando a los RFLP y a otros marcadores moleculares como las alozimas (Liu y Cordes, 2004; Chistiakov *et al.*, 2005). Su alta variabilidad, la simplicidad de sus protocolos de amplificación y las prácticas de muestreo poco invasivas

los han convertido en marcadores de elección para distintas aplicaciones tanto de la genética poblacional como de la genómica. Particularmente, el desarrollo de los nuevos marcadores de ADN ha contribuido al auge de los estudios de mapeo genético, haciendo asequibles estos proyectos incluso a especies de interés comercial menor. La Tabla I.2 resume las propiedades básicas de los marcadores de ADN más usados en los análisis de ligamiento genético, tanto en especies animales como vegetales.

Tabla I.2. Tipos de marcadores de DNA, características y algunas de sus principales aplicaciones.

Tipo de marcador	Información de secuencia <i>a priori</i>	Modo de herencia	Locus investigado	Polimorfismo	Principales aplicaciones
RFLP	Sí	Mendeliana, codominante	Único	Bajo	Mapeo de ligamiento, estudios poblacionales Mapeo de ligamiento, estudios poblacionales. Mapeo de ligamiento, estudios poblacionales Mapeo de ligamiento, estudios poblacionales Mapeo de ligamiento, estudios poblacionales.
RAPD	No	Mendeliana, dominante	Multiple	Intermedio	
AFLP	No	Mendeliana, dominante	Multiple	Alto	
SSR	Sí	Mendeliana, codominante	Único	Alto	
SNP	Sí	Mendeliana, codominante	Único	Alto	

Entre la variedad de marcadores basados en la PCR, en los últimos años los marcadores SNP se han revelado como los más útiles para el mapeo genético y la identificación de QTLs, dada su alta fiabilidad y su distribución a lo largo del genoma (1 cada 1300 pb en el genoma humano (Hartl y Jones, 1998; Sachidanandam *et al.*, 2001). Sin embargo, su caracterización implica un importante esfuerzo económico, ya que es necesario un desarrollo avanzado de la investigación genómica en la especie objeto de estudio, lo que en el ámbito de la acuicultura ha limitado su aplicación a las especies de mayor

interés comercial, como el pez gato (He *et al.*, 2003) y los salmónidos (Davey *et al.*, 2001), o científico, como las especies modelo, es el caso del pez cebra (Gjedrem, 2005) y medaka (Naruse *et al.*, 2000; Kasahara *et al.*, 2007). A lo largo de la última década, en la mayoría de los proyectos de mapeo de ligamiento en peces se han usado tres tipos de marcadores de ADN: microsatélites, RAPDs y AFLPs. Los tres tipos son marcadores clásicos en el mapeo de ligamiento tanto de especies animales como vegetales y si bien, una revisión bibliográfica reciente indica que la utilización de marcadores RAPD ha comenzado a decrecer, el uso de microsatélites y AFLPs estarían todavía en una fase exponencial de crecimiento (Liu y Cordes, 2004). En general, estos marcadores pueden clasificarse respectivamente como específicos de secuencia (microsatélites) e independientes de secuencia (AFLPs y RAPDs) (Danzmann y Gharbi, 2001).

1.3.2.2. Marcadores específicos de secuencia. Microsatélites.

Los marcadores específicos de secuencia requieren un conocimiento previo de las secuencias de ADN que flanquean al segmento de interés para el diseño de los cebadores o primers (Danzmann y Gharbi, 2001). Esta condición supone al mismo tiempo la principal ventaja y desventaja de este tipo de marcadores, puesto que incrementa su especificidad, pero también su coste. Marcadores de este tipo son los ya mencionados SNPs y los microsatélites.

Los microsatélites consisten en múltiples repeticiones dispuestas en tándem de una secuencia simple que puede variar en tamaño desde 1 a 6 pares de bases (pb) (Figura I.11; Tautz, 1989; Litt y Luty, 1989; Chistiakov *et al.*, 2005). Las formas más comunes de estos motivos de repetición son dinucleótidos tales como (CA)_n: (GT)_n, (GA)_n: (CT)_n, (CG)_n: (GC)_n, (AT)_n: (TA)_n, donde *n* es el número de repeticiones. En una serie en tándem, este número *n* es variable, pudiendo dar lugar a una gran cantidad de alelos para cada locus microsatélite en una población.

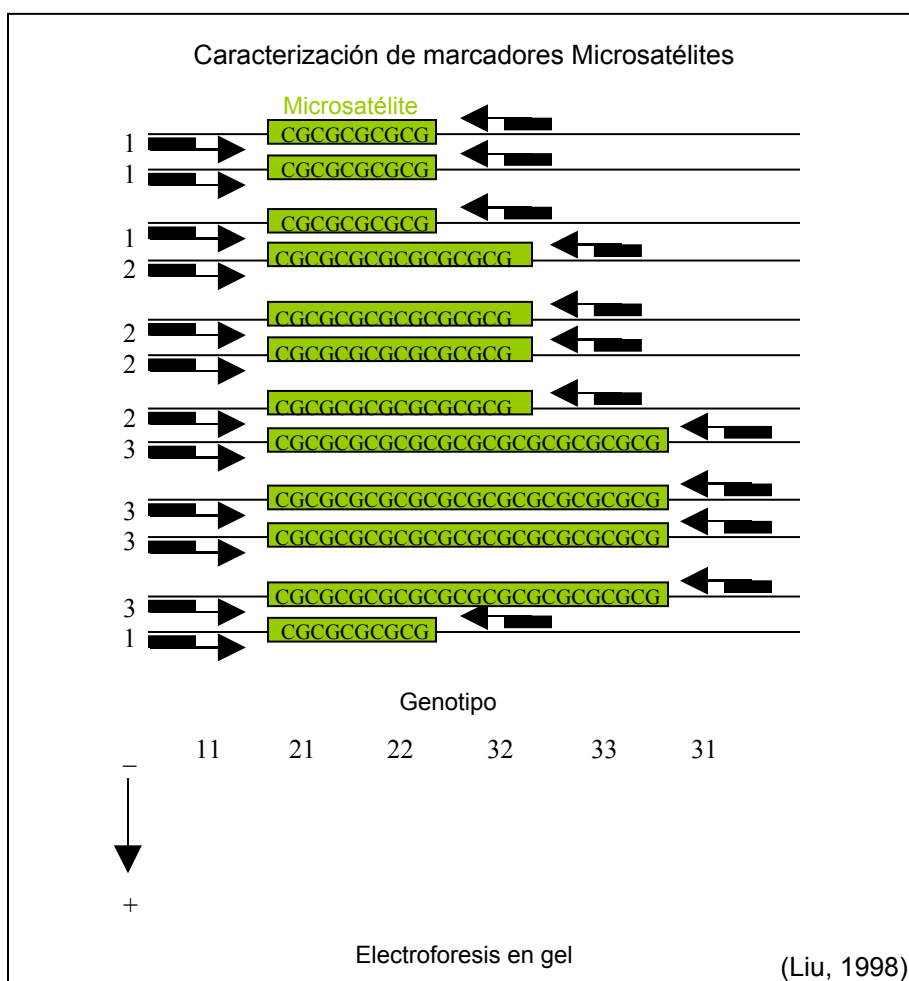


Figura I.11. Marcadores microsatélite. Principios de detección y fuente de polimorfismo. El polimorfismo microsatélite se basa en diferencias en el número de repeticiones en tándem, generalmente de motivos di-, tri- o tetranucleotídicos. En el ejemplo, se muestran los dos homólogos de un mismo brazo cromosómico para seis individuos. Para un único locus microsatélite $(CG)_n$ hay tres alelos diferentes dependiendo del valor de n . La variación en el tamaño de los alelos puede detectarse por PCR usando primers que flanquean la región de repetición. Para visualizar las diferencias en el número de pares de bases, los productos PCR se someten a electroforesis en un gel de separación. Para genotipos homocigóticos (11, 22 y 33) en el gel sólo se visualiza una banda, por el contrario, en los heterocigóticos (21, 32 y 31) la electroforesis muestra la expresión codominante de ambos alelos.

Además de su potencial polimórfico, los microsatélites reúnen una serie de características que los han convertido en marcadores de preferencia para el

mapeo génico: i) son numerosos y se distribuyen ampliamente a lo largo de todo el genoma, particularmente en peces se ha estimado que su frecuencia es de 1 cada 10 kb (Wright, 1993); ii) se heredan de modo mendeliano como marcadores codominantes, lo que facilita el seguimiento de su transmisión dentro de las poblaciones de mapeo (O'Connell y Wright, 1997); iii) son abundantes en todos los genomas eucariotas estudiados hasta la fecha (Hamada *et al.*, 1982; Tautz y Renz, 1984; Echt y Nelson, 1997; Liu y Cordes, 2004); iv) la mayoría de los loci microsatélite tienen un tamaño relativamente pequeño que facilita su caracterización vía PCR (Liu y Cordes, 2004); y v) son potencialmente transferibles entre linajes y/o especies próximas (Young *et al.*, 1998; Estoup y Angers, 1998), especialmente cuando el locus microsatélite se encuentra dentro de una secuencia génica (Pardo *et al.*, 2005; Nagy *et al.*, 2007). Esta última característica los hace especialmente útiles para estudios de genómica comparada, ya que convierte a los microsatélite en excelentes puntos de anclaje para identificar cromosomas homólogos entre especies próximas (Kondo *et al.*, 1993; Crawford *et al.*, 1995; McConnell *et al.*, 2000; Sakamoto *et al.*, 2000) o respecto a especies modelo bien conocidas a nivel genómico (Stemshorn *et al.*, 2005; Franch *et al.*, 2006).

A pesar de sus características favorables y de la aplicabilidad de los mapas construidos a partir de loci microsatélite (Johnson *et al.*, 1996; Shimoda *et al.*, 1999; Sakamoto *et al.*, 2000; Cristescu *et al.*, 2006), el uso de esta tecnología locus-específica resulta cara, especialmente cuando el objetivo último es la detección de QTLs, para lo cual se necesitan mapas densos, construidos a partir de cientos de marcadores que cubran adecuadamente el genoma. En especies poco caracterizadas a nivel genómico, se requieren estrategias alternativas, que permitan disponer de un número adecuado de marcadores a partir de unos recursos limitados. Este es el caso de muchas especies de peces, en las que en los últimos años se han llevado a cabo proyectos de mapeo que combinan el uso de marcadores microsatélite con marcadores multilocus de amplificación inespecífica como los AFLP, por

ejemplo, en la trucha arcoiris (*Oncorhynchus mykiss*, Nichols *et al.*, 2003), la platija japonesa (*Paralichthys olivaceus*, Coimbra *et al.*, 2003) y el salmón atlántico (*Salmo salar*, Moen *et al.*, 2004).

I.3.2.3. Marcadores independientes de secuencia. AFLP

Los marcadores anónimos o inespecíficos se desarrollan siguiendo metodologías que no requieren información previa acerca de la secuencia de ADN para la identificación y caracterización de los loci polimórficos. La PCR no se dirige a la amplificación específica de un locus de secuencia conocida, sino a la amplificación aleatoria de múltiples regiones repartidas a lo largo del genoma, cuyas secuencias flanqueantes resultan complementarias a cebadores diseñados al azar. Los primeros marcadores de este tipo fueron los RAPDs (Williams *et al.*, 1990). La identificación de este tipo de marcadores depende de la existencia de polimorfismos en los sitios de unión de los cebadores o bien de diferencias de longitud en los fragmentos amplificados. Las diferencias alélicas en un mismo locus se distinguen, generalmente, por la presencia o ausencia de una banda de un determinado tamaño.

El desarrollo de esta tecnología permitió la caracterización de numerosos marcadores en especies poco conocidas a nivel genómico, sin embargo, una de sus mayores limitaciones se refiere a la reproducibilidad. La amplificación de múltiples loci a partir de cebadores inespecíficos requiere utilizar condiciones PCR poco astringentes, lo que en ocasiones interfiere en la reproducibilidad de algunos marcadores (Ajmone-Marsan *et al.* 1997).

En 1995 Vos y colaboradores presentan la tecnología AFLP (*Amplified Fragment Length Polymorphism*), que combina las fuentes de variabilidad de RFLP y RAPD, al tiempo que supera sus principales limitaciones. El primer paso en la caracterización de marcadores AFLP es la digestión enzimática del ADN genómico de alto peso molecular. Como en los RFLP, esta digestión revelará polimorfismos basados en la inserción o delección de secuencias entre los sitios de corte o bien en la sustitución de bases en las propias dianas.

La detección de estos polimorfismos se lleva a cabo mediante la amplificación vía PCR. Como en los RAPD, esta amplificación sumará los polimorfismos derivados de la sustitución de bases en los lugares de unión de los cebadores. Lo novedoso de la tecnología AFLP es la incorporación entre las etapas de digestión y amplificación de un paso de ligamiento de adaptadores a los fragmentos resultantes de la restricción. Estos adaptadores de secuencia conocida servirán para el diseño de cebadores específicos que podrán unirse a los fragmentos en condiciones de amplificación muy astringentes, superando la baja reproducibilidad de los RAPDs (Ajmone-Marsan *et al.* 1997). Además, cada uno de estos cebadores incorpora una secuencia aleatoria de unas pocas bases, que distingue a cada pareja y permite la amplificación selectiva de un subconjunto de los fragmentos de restricción, lo que facilitará la interpretación de los patrones de bandas AFLP (Figura I.12).

Tanto RAPDs como AFLPs son esencialmente marcadores bialélicos dominantes, donde el genotipo homocigótico no puede diferenciarse del heterocigótico por la presencia de banda. Esta herencia dominante los convierte en marcadores menos informativos que los microsatélites o SNPs para la construcción de mapas. Los mapas basados en RAPDs o AFLPs son generalmente mapas preliminares densos que sirven de referencia para posteriormente posicionar otros tipos de marcadores menos numerosos, pero más informativos (Danzmann y Gharbi, 2001). Particularmente en peces, se han usado marcadores AFLPs y RAPDs para establecer mapas de referencia o “*framework*” en el pez cebra (Postlethwait *et al.*, 1994) y en la trucha arcoiris (Young *et al.*, 1998) o más recientemente en el pez gato (Liu *et al.*, 2003) y en el guppy o pez millón (Khoo *et al.*, 2003). En comparación con los RAPDs, los marcadores AFLP además de ser más reproducibles, han demostrado mayores niveles de polimorfismo y poder de diferenciación (Liu y Cordes, 2004). De hecho, en los últimos años, muchas publicaciones de mapas preliminares o de primera generación en especies de peces han usado preferentemente marcadores AFLP (Moen *et al.*, 2004; Poompuang y Na-Nakorn, 2004;

Watanabe *et al.*, 2004), especialmente, cuando no se dispone de los recursos para el desarrollo de marcadores microsatélite.

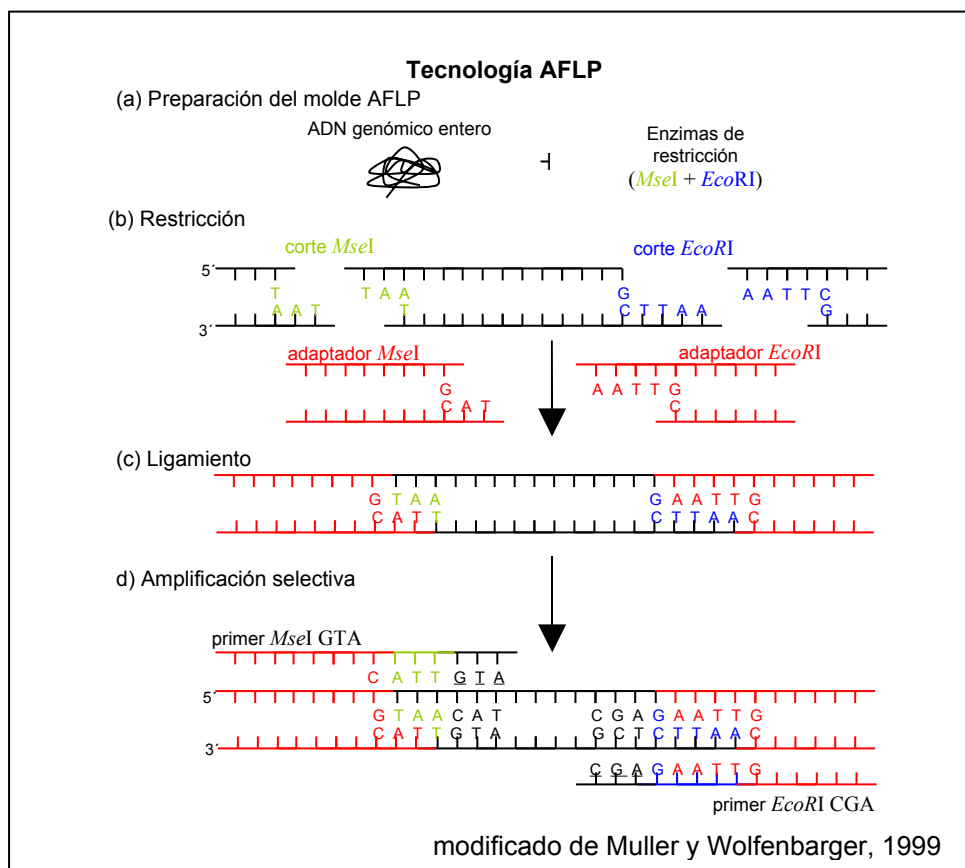


Figura 1.12. **Marcadores AFLP. Principios de detección y fuentes de polimorfismo.** La tecnología AFLP se basa en la amplificación selectiva vía PCR de un subconjunto de fragmentos de restricción (Vos *et al.*, 1995). Estos fragmentos se generan a partir de la digestión del ADN genómico de alto peso molecular por dos endonucleasas (a y b), una de corte menos frecuente (*EcoRI*) y otra más frecuente (*MseI*, en el ejemplo). A los fragmentos generados se unen adaptadores *Eco* y *MseI* (c) a fin de crear lugares de unión para los cebadores en la amplificación por PCR (d). El uso de nucleótidos selectivos en los extremos 3' de los primers AFLP permite una amplificación selectiva de un subconjunto de los fragmentos de restricción. La migración electroforética de los productos PCR revela un perfil multilocus. La presencia de polimorfismos en los sitios de restricción o en el acoplamiento selectivo de los cebadores, lleva a la identificación de los marcadores AFLP dominantes como presencia o ausencia de fragmentos de un determinado tamaño molecular en los perfiles de migración.

I.3.3. Análisis computacional del ligamiento genético

El genotipado de marcadores en la población de mapeo genera un conjunto de datos bruto a partir del cual se interpretarán las relaciones de ligamiento entre loci. La detección de ligamiento y la consiguiente construcción del mapa procede fundamentalmente en tres etapas: análisis de locus-único o comprobación de la herencia mendeliana; análisis de dos-puntos o test de asociación entre pares de marcadores; y análisis multi-punto u ordenación de marcadores.

I.3.3.1. Análisis de los loci individuales

El análisis de loci únicos se refiere a la estima de las frecuencias alélicas y a la comprobación del ajuste de la segregación a las proporciones mendelianas esperadas según el modelo genético utilizado.

El modelo de herencia depende de la población de mapeo seleccionada para la construcción del mapa. Por ejemplo, en un retrocruzamiento, los genotipos heterocigótico y homocigótico para cada marcador deben encontrarse en proporción 1:1 en la progenie. Los marcadores que se ajustan a las proporciones previstas se seleccionan para la construcción del mapa, sin embargo, también aquellos que demuestran desviaciones significativas de las proporciones de segregación esperadas pueden contener información valiosa para el mapeo.

La distorsión en la segregación de los marcadores puede deberse a un muestreo inadecuado de los individuos o a errores de genotipado relacionados con problemas técnicos o humanos, sin embargo, detrás de la desviación pueden estar también factores biológicos. La detección de desviaciones significativas de las proporciones de segregación previstas podría estar identificando a marcadores asociados a genes sujetos a selección, por lo que la inclusión de estos marcadores en el análisis de ligamiento puede aportar una valiosa información al mapa.

I.3.3.2. Análisis de dos puntos.

Una vez verificada la calidad de los datos genotípicos, la elaboración del mapa genético comienza con el análisis de la asociación entre marcadores. La magnitud de datos conlleva el uso de programas informáticos específicamente diseñados para llevar a cabo los cálculos computacionales implicados en la determinación de la asociación, orden y distancias entre los marcadores ligados. Inicialmente, los programas proceden calculando las frecuencias de recombinación entre todos los posibles pares de marcadores y la detección del ligamiento entre ellos, por lo que se conoce a esta etapa como análisis de dos-puntos.

Generalmente, para estimar la frecuencia de recombinación entre dos marcadores se utiliza el método de máxima verosimilitud (Dempster *et al.*, 1977; Lander *et al.*, 1987; Morton, 1995; Ott, 1999). El valor de recombinación (r) resultante se testa entonces frente a la hipótesis nula de no asociación ($r = 0.50$) a fin de determinar si existe ligamiento entre el par de marcadores. En casi todos los análisis de ligamiento actuales, el test de asociación se resuelve usando el Método Lod Score o LOD, es decir, el logaritmo en base 10 de la razón entre el valor de la función de máxima verosimilitud para r y el valor de la función para segregación independiente (Morton, 1995). Si la razón de verosimilitudes es mayor que un límite predeterminado, se rechazará la hipótesis nula, y se considerará que los marcadores están ligados. Dado que la probabilidad *a priori* de que dos marcadores elegidos al azar estén ligados es muy baja, usando cálculos bayesianos se estimó que sería necesaria una evidencia de 1000:1 a favor de ligamiento para obtener el umbral de significación convencional de $P \leq 0,05$ (Strachan y Read, 1999).

Finalmente, si el locus A resulta ligado al locus B, y B está ligado a C, los tres marcadores formarán parte de un mismo grupo de ligamiento. El programa informático procede con esta lógica analizando todos los pares de marcadores y distribuyéndolos en grupos de ligamiento, que en los mapas más

densos se corresponderán con el número haploide de cromosomas de la especie.

I.3.3.3. Análisis multipunto.

Una vez que se ha identificado los grupos de ligamiento, el siguiente paso en la construcción de un mapa es averiguar el orden lineal de los marcadores a lo largo de cada grupo y calcular las distancias entre los loci ordenados.

Cuando hay loci no informativos en una fracción significativa de las meiosis, como ocurre con frecuencia en los pedigrís de mapeo, el análisis de tres puntos no es suficiente para establecer entre qué dos marcadores es más probable que se localice un tercero. Se hace necesario entonces el análisis simultáneo de múltiples loci a fin de asegurar que en todas las meiosis estén presentes marcadores flanqueantes informativos. Así, para ordenar los marcadores dentro de un grupo de ligamiento, los programas de mapeo incluyen algoritmos para el análisis simultáneo de múltiples loci (Weir 1996, Ott 1999; Nuez 2000). Este análisis multipunto combina dos aspectos complementarios: la estimación de la frecuencia de recombinación multilocus entre marcadores seleccionados y la determinación del orden de marcadores más probable dentro de un grupo de ligamiento usando un criterio de optimización basado en identificar el orden de marcadores que minimice el número de entrecruzamientos.

Las estimas de frecuencia de recombinación multipunto podrían ser suficientes para calcular las distancias de mapa si los sucesos de recombinación fueran independientes unos de otros. Sin embargo, en el cálculo de las distancias ha de tenerse en cuenta el efecto de la interferencia cromosómica. Considerando éste fenómeno, los programas de mapeo incluyen funciones de mapa, generalmente las de Haldane y Kosambi, para la conversión de las frecuencias de recombinación en distancias genéticas aditivas. La elección de una u otra función de mapa para el cálculo de

distancias dependerá fundamentalmente del patrón de interferencia considerado en la especie en estudio.

I.4. MAPEO GENÓMICO EN PECES.

El término “pez” comprende un grupo complejo de vertebrados, desde los primitivos Agnata a los Teleostei (Nelson, 1994), representando una amplia variación de tamaños genómicos desde las 400 Mb en algunos Tetraodontidae hasta las 100.000 Mb en los peces pulmonados africanos (Hinegardner, 1968; Hinegardner y Rosen, 1972; Ohno, 1974; Tiersch *et al.*, 1989). Dentro de un grupo tan diverso como éste de los peces, el valor económico y/o científico de ciertas especies han dirigido los esfuerzos en investigación genómica, ampliando el conocimiento genético en especies puntuales, mientras que la mayoría de las especies de peces siguen siendo grandes desconocidas a nivel molecular. Particularmente, tres especies: *Danio rerio* (pez cebra), *Takifugu rubripes* (fugu) y *Tetraodon nigroviridis*, son en la actualidad objeto de intensos programas de secuenciación, debido a su valor como organismos modelo. Fugu y Tetraodon se han usado como modelo para comparaciones genómicas con humanos (Roest Crollius *et al.*, 2000; Rothenberg, 2001). El pequeño tamaño del genoma de fugu motivó que se convirtiera en el segundo vertebrado en ser secuenciado. Sin embargo, ninguno de ellos cría en cautividad por lo que, el desarrollo de líneas consanguíneas para la elaboración de mapas de ligamiento está excluido. Por el contrario, el pez cebra, bien conocido por su importancia como organismo modelo en biología del desarrollo y de gran interés acuarístico, cría fácilmente en cautividad (Nusslein-Volhard, 1994) y es, actualmente, una de las pocas especies de peces para la que se dispone de un mapa genético de alta densidad (Shimoda *et al.*, 1999).

Además del interés científico, el desarrollo de la industria piscícola, que genera importantes ganancias económicas, ha motivado un creciente interés en la construcción de mapas de ligamiento en especies comerciales como los

ya citados anteriormente del salmón atlántico (*Salmo salar*), la lubina (*Dicentrarchus labrax*), la dorada (*Sparus aurata*) o la trucha arcoiris (*Oncorhynchus mykiss*), principales especies en la industria acuícola europea (Clark, 2003). La construcción de estos mapas se dirige fundamentalmente a la identificación de QTLs relacionados con caracteres de interés económico, tales como la tasa de crecimiento, la resistencia a patógenos, la tolerancia a determinados niveles de salinidad, etc. El mapa se convierte entonces en una herramienta de notable valor para la mejora de la producción industrial.

En las especies acuáticas, la construcción de los primeros mapas se veía limitada por la escasez de marcadores, restringidos a la información obtenida de líneas fenotípicas mutantes y a unos pocos marcadores bioquímicos (revisión en Lie *et al.*, 1994). Sin embargo, el avance en el desarrollo de marcadores basados en la reacción en cadena de la polimerasa (PCR) como RAPDs, AFLPs, microsatélites y más recientemente, los SNPs, repercutió también en el conocimiento de los genomas de organismos acuáticos, facilitando la construcción de mapas de ligamiento genético más densos. Desarrollada la tecnología, la utilización de un tipo u otro de marcador depende entonces de la información de secuencia de la que se disponga en cada especie. De hecho, pocos mapas genéticos en peces incluyen marcadores de secuencias génicas (Sakamoto *et al.*, 2000; Lee *et al.*, 2005; Gharbi *et al.*, 2006; Wang *et al.*, 2007).

Por otro lado, la estrategia utilizada para seguir la segregación de estos marcadores en las familias de referencia, viene condicionada por la disponibilidad de líneas consanguíneas en la especie objeto de mapeo. En especies de peces donde se dispone de líneas consanguíneas, la población de mapeo generalmente consiste en una F_2 (Khoo *et al.*, 2003; Kimura *et al.*, 2005; Lee *et al.*, 2005) o en un retrocruzamiento de tipo 1, es decir, $F_1 \times P$ (Liu *et al.*, 2003; Watanabe *et al.*, 2004) o de tipo 2, $F_2 \times P$ (Gilbey *et al.*, 2004). “P” indica a cualquiera de los progenitores, que pertenecen a líneas puras y

generalmente se eligen los más diferenciados a nivel genético, a fin de maximizar la variabilidad. En algunos casos, para mejorar la información obtenida de marcadores dominantes, la población de mapeo la constituye un único parental F_1 procedente de un cruce entre líneas consanguíneas o entre especies afines y su descendencia ginogenética o androgenética (Young *et al.*, 1998; Nichols *et al.*, 2003).

En especies salvajes o en aquellas que empiezan a ser domesticadas y para las cuales los programas de selección genética están en las primeras etapas no es posible disponer de líneas consanguíneas y es necesario buscar estrategias alternativas para seguir la segregación de los marcadores desde los parentales a la descendencia. Se recurre en estos casos al diseño de pseudo-retrocruzamientos (Wilson *et al.*, 2002; Chistiakov *et al.*, 2004; Kai *et al.*, 2005) y poblaciones haploides o diploides procedentes de un único parental cuando la tecnología de ginogénesis o androgénesis está a punto (Kocher *et al.*, 1998; Poompuang y Na-Nakorn, 2004).

Los esfuerzos en el mapeo genético de peces han llevado a incrementar el conocimiento de la organización del genoma del grupo más diverso de Vertebrados. A medida que aumente el número de mapas disponibles, podrá estudiarse con mayor profundidad el nivel de conservación del orden y la sintenia de los principales genes ortólogos entre las especies de peces, proporcionando los datos necesarios para reconocer los reordenamientos cromosómicos que hayan tenido lugar a lo largo de su evolución y ayudando así a entender su complejidad (Schoen, 2000; Danzmann y Gharbi, 2001).

I.5. OBJETIVOS

El principal objetivo del presente trabajo es la construcción de un primer mapa de ligamiento genético en rodaballo (*Scophthalmus maximus*), que contribuya al conocimiento genómico de esta especie. Con este fin, se desarrollarán un conjunto de marcadores genéticos, que se sumarán a los ya disponibles en esta especie para lograr una densidad mínima que permita caracterizar sus relaciones de ligamiento y obtener un mapa preliminar. Este mapa sentará las bases para el desarrollo de mapas de segunda generación, en los que poder identificar loci implicados en caracteres cuantitativos (QTLs) y/o en resistencia a enfermedades. La elaboración de este primer mapa se puede resumir en tres pasos que constituyen, a su vez, los objetivos específicos de este trabajo:

- (i) Obtención de una familia de referencia para el mapeo genético. Teniendo en cuenta el desarrollo de las técnicas de manipulación cromosómica en esta especie, se utilizarán familias de ginogenéticos haploides para el mapeo, dada la naturaleza mayoritariamente dominante de los marcadores usados en este estudio.
- (ii) Puesta a punto de nuevos marcadores moleculares en rodaballo. Dada la escasez de marcadores disponibles en esta especie, su desarrollo y caracterización será uno de los objetivos esenciales de este trabajo. Se aplicará por primera vez en esta especie la tecnología AFLP, que no requiere un conocimiento genómico previo para la identificación de marcadores y a juzgar por la literatura, muestra una consistencia técnica apropiada para la tarea que se pretende abordar. Se incluirá también la puesta a punto de marcadores microsatélite, codominantes, polimórficos y muy consistentes. Sin embargo, la presencia de estos marcadores será

menor que la de los AFLPs, puesto que su desarrollo supone una importante inversión tanto económica, como de tiempo.

- (iii) Construcción del mapa de ligamiento. El tercer y último objetivo supone la identificación de las relaciones de ligamiento entre los marcadores y la ordenación de los mismos en función de sus frecuencias de recombinación. El cálculo de las distancia de mapa entre los marcadores ordenados, permitirá obtener una estima del tamaño genómico en cM de esta especie.

I.MATERIAL Y MÉTODOS

II. 1. MATERIAL BIOLÓGICO

El primer paso en un proyecto de mapeo es la selección o en su caso, el desarrollo, de una población de referencia que permita seguir la segregación alélica de los marcadores desde los parentales a su descendencia. En el caso del rodaballo, no se dispone de líneas puras a partir de las cuales se puedan diseñar cruzamientos prueba adecuados a la estrategia de mapeo. Sin embargo, en esta especie se ha aplicado con éxito la tecnología para la inducción de ginogénesis, obteniendo progenies haploides genéticamente idénticas a la madre (Piferrer *et al.*, 2004). La transmisión génica en estas familias se comporta a efectos de segregación como un retrocruzamiento, lo cual las convierte en paneles de mapeo ideales para el análisis de la herencia tanto de marcadores codominantes como dominantes.

En esta memoria, la familia de referencia para el desarrollo del mapa genético se constituyó a partir de una hembra adulta procedente de una población natural de rodaballo atlántico y su progenie de 50 embriones ginogenéticos haploides. Trabajar con una familia de estas características implica una serie de particularidades a nivel metodológico en cuanto a la inducción de ginogénesis y a la extracción de ADN de los embriones haploides resultantes. Superadas estas etapas y verificada la contribución exclusivamente materna a la descendencia, la selección de la familia de referencia se basó en la heterozigosis de la hembra progenitora, como estimador de su informatividad para la construcción del mapa. Cada una de estas etapas se detallan en los siguientes apartados.

II.1.1. Inducción de ginogénesis

En especies piscícolas, la mayor tasa de crecimiento de las hembras bajo condiciones de cautividad (Imsland *et al.*, 1997) ha llevado a invertir esfuerzos económicos y de investigación dirigidos a la obtención de poblaciones femeninas. Como resultado, existen en la actualidad diferentes

métodos para la inducción de ginogénesis en especies de interés comercial (Piferrer *et al.*, 2000; 2004; Felip *et al.*, 2001). Los más usuales son los métodos basados en la fertilización de los huevos con espermatozoides que han sido genéticamente inactivados, bien por exposición a radiación ultravioleta o por aplicación de tratamientos químicos (Thorgaard, 1983; Ihssen *et al.*, 1990; Felip *et al.*, 2001). Este espermatozoides alterado debe mantener la movilidad, de forma que pueda activar el desarrollo embrionario sin hacer ninguna aportación genética.

La familia de rodaballos ginogenéticos haploides utilizada en esta tesis se obtuvo en el Instituto Español de Oceanografía (I.E.O) de Vigo, por fertilización de huevos con espermatozoides de rodaballo previamente irradiados con dosis de 30,000 erg mm⁻² de luz UV, de acuerdo al método descrito por Piferrer *et al.* (2004). La dosis de radiación UV necesaria para inactivar el ADN sin comprometer la capacidad motriz del espermatozoides se determinó comprobando el tiempo de exposición a la radiación en diferentes fracciones de la puesta de la hembra.

Debido a la haploidía, los embriones resultantes sufren un desarrollo anormal y mueren antes de la eclosión, fenómeno conocido como “síndrome haploide” (Lie *et al.*, 1994). La edad de los embriones haploides de rodaballo fue determinada en el I.E.O y en ningún caso superó las 120 h de vida.

II.1.2. Extracción de ADN a partir de familias ginogenéticas haploides

La calidad del ADN es un factor fundamental a la hora de genotipar correctamente una muestra de individuos. La calidad, así como la cantidad del ADN disponible van a depender del procedimiento de extracción que, a su vez, estará condicionado por la naturaleza de las muestras, así como por el estado de conservación de los tejidos de partida.

El ADN se aisló en la familia de referencia siguiendo un protocolo de fenol: cloroformo: alcohol isoamílico. Los individuos haploides presentan, sin embargo, una particularidad inherente a su propia condición que complica el proceso de extracción, la escasa cantidad de ADN disponible. Por definición, la cantidad de ADN de un individuo haploide en un determinado estadio de desarrollo es la mitad de la cantidad que correspondería a un individuo diploide de la misma edad. Por otro lado, los efectos del ya mencionado síndrome haploide condicionan el muestreo a individuos en estado embrionario con un limitado número de células. Atendiendo a estas particularidades, el protocolo de extracción hubo de ser modificado para obtener, a partir de embriones haploides, ADN de alto peso molecular adecuado para la aplicación de metodologías multilocus (AFLP) y en cantidad suficiente para llevar a cabo la genotipación de más de 100 marcadores de ADN por individuo.

II.1.2.1. Extracción de ADN genómico de alto peso molecular en rodaballos diploides adultos.

En los progenitores, la extracción de ADN se realizó a partir de tejido de aleta siguiendo el procedimiento clásico fenol: cloroformo: alcohol isoamílico (Sambrock *et al.*, 1989; Apéndice I).

II.1.2.2. Extracción de ADN genómico de alto peso molecular en embriones haploides de rodaballo.

Las principales modificaciones con respecto al protocolo clásico de fenol: cloroformo: alcohol isoamílico se refieren a cuatro aspectos:

i) La proporción entre la concentración de proteinasa K y la cantidad de tejido. Algunos protocolos de extracción de ADN genómico a partir de embriones describen la eliminación manual del saco vitelino (Kocher *et al.*, 1998; Poompuang y Na-Nakorn, 2004; Sun y Liang, 2004). Sin embargo, el escaso tamaño de los embriones de rodaballo no permite la separación física de saco y embrión y ha de incrementarse la cantidad de proteinasa K para conseguir la digestión del tejido.

ii) Utilización de una barrera inerte en el proceso de separación de las fases acuosa y orgánica (Phase Lock Gel (PLG), Eppendorf), que actúa evitando el contacto entre ambas fases, aumentando así la eficacia de la extracción de ADN.

iii) Incremento del tiempo de precipitación del ADN en etanol y reducción del número de lavados.

El protocolo final se detalla en el apéndice II.

II.1.3. Selección de la familia de referencia

La familia de referencia para el análisis de ligamiento se seleccionó entre 14 familias de rodaballo sometidas al protocolo para inducción de ginogénesis en el I.E.O (Tabla II.1).

La confirmación del origen exclusivamente materno de los embriones se basó en el genotipado de marcadores microsatélite diagnóstico en 20 descendientes de cada una de las familias supuestamente ginogenéticas (para detalles de la amplificación PCR consultar apartado II.2.1 y apéndice IV). Se consideraron marcadores diagnóstico, aquellos microsatélites para los cuales, los dos progenitores de un determinado cruce portan 4 alelos diferentes. Igualmente, en los casos dónde los óvulos de una única hembra fueron activados por esperma de un pool de machos (familias H85, H47 y H20), se comprobó que ninguno de los alelos de los machos implicados coincidiera con los alelos de la hembra. En todos los casos, la diferencia de tamaño entre los alelos de cada progenitor hubo de ser de al menos dos pares de bases, de forma que pudieran distinguirse sin ambigüedad en un análisis de fragmentos.

Tabla II.1. Para cada familia analizada se indican su código y el cruce que la origina (H: hembra; M: machos), los loci microsatélite seleccionados como diagnóstico para chequear la herencia exclusiva materna y los correspondientes genotipos materno y paterno. En algunos casos, se utilizó el esperma de varios machos para activar huevos de una única hembra.

Familia	cruce	Locus	Genotipo materno	Genotipo paterno
H8	H8xM40	Smax-02 Sma1-125INRA	95/129 120/136	111/117 132/132
H104	H104 x M379	Smax-04	78/116	118/122
H37	H337 x M125	Smax-04 Sma3-129INRA	118/122 175/187	78/116 179/185
H279	H279 x M79	Smax-02 Sma3-12INRA	111/121 88/108	101/127 92/100
H226	H226 x M578	Smax-04 Sma3-129INRA	90/90 189/191	78/122 175/179
H768	H768 x M019	Sma3-129INRA	178/185	185/187
H68	H68 x M14	Smax-04 Sma1-125INRA	90/118 136/136	116/120 140/152
H17	H17 x M20	Sma4-14INRA Sma1-125INRA	134/136 112/132	124/130 136/136
H85	H85 x M75 M40 M21	Smax-04 Sma1-125INRA	112/120 120/136	116/118 90/116 116/118 132/132 132/132 132/132
H67	H67 x M23	Sma1-125INRA	128/132	136/136
H06	H06 x M21	Sma3-12INRA Sma1-125INRA	92/96 140/140	88/94 132/132
H18	H18 x M60	Sma4-14INRA	124/130	126/126
H47	H47 x M30 M10	Sma3-8INRA Sma1-125INRA	186/202 132/136	192/198 192/208 112/128 120/128
H20	H20 x M79 M60	Smax-02 Sma3-12INRA	95/95 181/191	101/127 107/111 187/189 179/187

Los progenitores de las 14 familias formaban parte de un stock de rodaballo atlántico del I.E.O que anteriormente había sido genotipado para un total de 13 marcadores microsatélite (Coughlan *et al.*, 1996; Estoup *et al.*, 1998; Bouza *et al.*, 2002). Esta información facilitó la selección de marcadores diagnóstico. Según Castro *et al* (2003) el análisis de dos loci diagnóstico en la progenie es suficiente para confirmar la contribución exclusivamente materna al genoma de la descendencia. En este trabajo se siguió esta recomendación, con la excepción de cuatro cruces (H104 x M379; H768 x M019; H67 x M23; y

H18 x M60) en los que sólo se encontró un locus microsatélite diagnóstico, entre los 13 caracterizados. Las familias que presentaron en la descendencia alguno de los alelos paternos fueron descartadas.

En cada familia ginogenética, la madre se genotipó para otros 18 marcadores microsatélite de rodaballo (Iyengar *et al.*, 2000) más cinco nuevos microsatélite, Sma-USC1, Sma-USC2, Sma-USC3, Sma-USC4 y Sma-USC5, caracterizados a lo largo del desarrollo de esta memoria (para detalles consultar el apartado II.2.1.2 y apéndice IV). Estos nuevos genotipos se sumaron a los de los 13 microsatélites ya disponibles. Finalmente, la familia de referencia para la construcción de mapa, se seleccionó de acuerdo al nivel de heterozigosis de la hembra y al tamaño de la progenie.

II.1.4. Amplificación genómica global: GenomiPhi™.

A pesar de la optimización del protocolo de extracción, la cantidad de ADN obtenida a partir de los embriones haploides resultaba escasa para la caracterización del número de marcadores requerido en un proyecto de mapeo genómico, aún siendo de moderada densidad. A fin de superar esta limitación, en la familia de referencia seleccionada se aplicó un protocolo de amplificación genómica (GenomiPhi™; Amershan Bioscience) al ADN extraído de sus embriones haploides.

El kit de amplificación genómica GenomiPhi™ de Amershan Bioscience usa la ADN polimerasa del bacteriófago *Phi29* para amplificar exponencialmente moléculas de ADN lineal. El proceso transcurre en una reacción isotérmica (30°C) durante la cual el enzima puede generar cantidades de ADN del orden de microgramos a partir de nanogramos de la molécula inicial.

El método de amplificación se basa en el emparejamiento aleatorio de cebadores hexaméricos a lo largo del ADN genómico (Figura II.1). De esta forma, la polimerasa *Phi29* inicia simultáneamente el proceso de replicación en múltiples sitios a lo largo del ADN molde. La particularidad de este enzima es que su actividad no se detiene al encontrar las nuevas cadenas de ADN en la dirección de síntesis, sino que es capaz de desplazarlas y continuar su proceso de replicación. La unión de cebadores adicionales a las nuevas moléculas de ADN de cadena simple resulta en la producción de grandes cantidades de ADN de cadena doble de alto peso molecular.

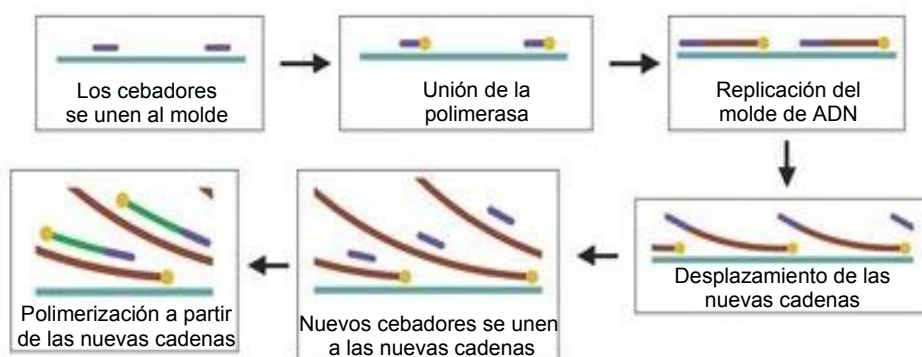


Figura II.1. Esquema del proceso de amplificación con el método GenomiPhi™. Los cebadores hexaméricos se unen en múltiples puntos al ADN molde desnaturizado. La Polimerasa *Phi29* inicia la replicación simultáneamente en cada punto. A medida que ocurre la síntesis de ADN, las nuevas cadenas van desplazando a las amplificaciones contiguas también de nueva síntesis. La subsiguiente replicación de las nuevas cadenas lleva a la formación de ADN de cadena doble (modificado de Amershan Bioscience).

Para evitar contaminaciones, el protocolo de amplificación se llevó a cabo en una cámara de flujo previamente esterilizada por tratamiento con luz UV tal como se detalla en apéndice III. El éxito de la amplificación genómica se comprobó por electroforesis en un gel de agarosa al 1%. A partir de cada muestra amplificada se preparon alícuotas de 10 ng/μl en un volumen total de 100 μl.

II.2. DESARROLLO Y CARACTERIZACIÓN DE MARCADORES MOLECULARES.

Un factor limitante a la hora de abordar un proyecto de mapeo es la disponibilidad de marcadores de ADN en la especie objeto de estudio. En el caso de *S. maximus*, al comienzo de este trabajo, tan sólo se encontraban en la literatura 31 marcadores microsatélite para esta especie (Coughlan *et al.*, 1996; Estoup *et al.*, 1998; Iyengar *et al.*, 2000; Bouza *et al.*, 2002). A fin de incrementar el número de marcadores disponibles para el mapeo se siguieron dos estrategias diferentes: identificación y puesta a punto de nuevos marcadores microsatélite y aplicación de la tecnología AFLP (*Amplified Fragment Length Polymorphism*, Vos *et al.*, 1995).

II.2.1. Marcadores microsatélite

La Tabla II.2 incluye los 31 marcadores microsatélite de rodaballo disponibles en la literatura en el momento de abordar este trabajo de Tesis.

Tabla II.2. Secuencias de los cebadores y motivos de repetición de los 31 loci microsatélite publicados en rodaballo. Referencias bibliográficas: (1) Coughlan *et al.*, 1996, (2) Bouza *et al.*, 2002, (3) Estoup *et al.*, 1998, (4) Iyengar *et al.*, 2000.

Locus	Nº de acceso	Secuencias de los cebadores (5'-3')	Motivo de repetición	Ref.
Smax-01	Z78097	GGTGATAACGAGTTTGATGA TTCACAAGTAATCTCCCAAC	(TG) ₂₀	1
Smax-02	Z78101	GGAGGATGTATTGAAAGTGT AGAGCAGGTCATTATACAGC	(TG) ₁₆	1
Smax-03	Z78098	TGGTTAATTACAGCCTTTGG CTCATTTCTTTTTGTTTGC	(AC) ₁₆	1
Smax-04	Z78099	TCACAAAATCTGAACATCGGT GTAATCCACCAAAGTGAACG	(TG) ₂₀	1
Smax-04b	Z78100	GGGTCATCATCTTTAGTCT GTCACAGTTTCATTAGGAG	(AC) ₂₃	1,2
Sma3-8INRA	-	CCCTCCGTCAGACAAAGAG GACGAAGTTAATTTTCATTG	(GT) ₁₈ -(CT) ₅	3
Sma3-10INRA	-	CTATGCAAGTGAAGACACGTAC CTTTCACCTCTGTTGTATCTAAGG	(TG) ₅₀	3
Sma3-12INRA	-	CACAATTGAATCACGAGATG GCCACCACTCGCTAACAC	(TG) ₂₁	3
Sma4-14INRA	-	GCGAGCAAATATCAGAAGG CCAGAAACAGCTCCGACTC	(GT) ₂₁	3
Sma5-111INRA	-	TCTACACTGCAGGTTGGG CTGATTATGGGCTGGACG	(TG) ₁₃	3

Tabla II.2 Continuación

Locus	Nº de acceso	Secuencias de los cebadores (5'-3')	Motivo de repetición	Ref.
Sma1-125INRA	-	CACACCTGACAAAGCTCAAC GCTGAACATTTTCATGTTGATAG	(TAGA) ₁₁ -(TG) ₄	3
Sma3-129INRA	-	GCACTGCCTTTTCATTGG CAGCTCTAGATTGTTTATCCC	(GT) ₂₉	3
Sma1-152INRA	-	GGGGAGGTGTCTGTCTGG GCACAGCTCCGAGCACAG	(GT) ₁₄	3
B11-I 12/6/3	AF182080	GTTGTTACTGCTGATCTAGC ATGTTCCATCTCATTCCCTGC	(CA) ₁₂ CG (CA) ₆ CG(CA) ₃	4
B12-I GT14	AF182082	GTGATGGAAGATTGTACCAG CACAATAAAGGATAGACAGG	(GT) ₁₄	4
B18-II CA70	AF182083	CAGAGGGGATAATTCTGTGC TACTGGCATCATGGTCAAC	(CA) ₇₀	4
F1-OCA19	AF182084	AGTTACACCAAGTGCACAGAG CCAGGCCATCCACATTTAAC	(CA) ₁₉	4
F8-ICA26	AF182085	ATTCCTACGAGTCAGGCTCG TTCGATGTCAGTCGAGGACG	(CA) ₂₆ (GA) ₃	4
F8-I 11/8/17	AF182086	CACTAATTACACGGCAGACG AATCGGATGAGACGGCGAAC	(CA) ₁₁ T(CA) ₈ T(CA) ₁₇	4
F12-I TG	AF182087	CAGTCATATCAGCAGATGCC AATTCAGCCGACAACTCTC	(TG) ₁₆ (CG) ₄	4
F12-I AG18	AF182010	GGCAGGACTGTACATTTTGG AGTCAGCCAGTTTACTGCAC	(AG) ₁₈ G (GA) ₁₅	4
1/4AC18	AF182088	AGAAAGGCTCGACCAGCTCC TGATGGCTCATAGTGGCTAC	(AC) ₁₈	4
2/5TG14	AF182089	GAGGATTGGTGACCTTGGC CCTCAGACTACTGAAGTCAC	(TG) ₁₄	4
3/3GT	AF182090	GATCTAGTTGTCTAACAGG ATAGCAGGTCCTGCTAAAG	(GT) ₁₄ GC (GT) ₈ (TG) ₄ T(TG) ₃	4
3/9CA15	AF182091	AGAGTGAAGAACGTACCTGC CAATGGAGAGGCAGTATCGG	(CA) ₁₅	4
3/20CA1	AF182092	AAGCACCCATACGCAGG AAGCACCCATACGCAGG	(CA) ₁₇	4
4/3COMP	AF182093	GTACTTAAGCAGCTGTTTCGC GAAACGCAGTTAGTTCCTGC	(A) ₅ CA (AC) ₄ CA(AC) ₄ G(AC) ₅ G(AC) ₇ ATG(AC) ₂₀ ATT (CA) ₁₃ AT(CA) ₃ (AC) ₄ G(CA) ₁₃	4
4/4CA4/13	AF182094	GCACTCATGCTGACATACAG GACTCAGTGACTGATGAATC	(AC) ₄ G(CA) ₁₃	4
4/5CA22/6 /2	AF182095	ACTCTGAGGGCAGAGGAAC GCTGCTGGTAATCAGAAAGC	(CA) ₂₂ CT(CA) ₆ CT(CA) ₂	4

Tabla II.2 Continuación

Locus	Nº de acceso	Secuencias de los cebadores (5'-3')	Motivo de repetición	Ref.
5/4CA20	AF182096	GACGGAAGCTCGGAGATTC GATTGGTGACCTTGGCAAC	(CA) ₂₀ A(AC) ₂	4
7/1TC18	AF182097	GAGAACTGAGCTGAGTCAC AATGATTATCACCGCTGCAC	(TC) ₁₈	4

II.2.1.1. Estrategias para la obtención de marcadores microsatélite

Para la identificación de nuevos marcadores microsatélite en rodaballo, se siguieron dos estrategias distintas: 1) amplificación en rodaballo de loci microsatélite descritos en especies próximas; y 2) aislamiento de loci microsatélite a partir de una librería genómica enriquecida en motivos de repetición cortos en tándem (STR, *Single Tandem Repeat*).

II.2.1.1.1. Amplificación cruzada.

En total, se testaron en rodaballo 17 parejas de cebadores de marcadores microsatélite polimórficos procedentes de dos especies de peces planos: 7 caracterizados originalmente en la platija japonesa (*Pleuronectes platessa*, Hoarau *et al.*, 2002) y los otros 10 en el fletán (*Hippoglossus hippoglossus*, Coughlan *et al.*, 2000) (Tabla II.3). La amplificación de cada par de cebadores se testó en dos individuos de una población natural de rodaballo del Cantábrico.

Tabla II.3. Marcadores microsatélite de platija (*P. platessa*) y fletán atlántico (*H. hippoglossus*) testados en rodaballo. Referencias bibliográficas: (1) Hoarau *et al.*, 2002, (2) Coughlan *et al.*, 2000.

Locus	Nº de acceso	Secuencia del cebador (5'-3')	Motivo de repetición	Ref
PL06	AF406745	AACTGATTTACTTTACTTTGTC GTCGTCCTCCAAATTGAAC	(GT) ₇ TT(GT) ₂₅	1
PL09	AF406746	GTAATTTTCTAACATATGGAGG TCACACCCAAGGCTACG	(CA) ₄₉	1
PL52	AF406747	GAAACCCCTGGATCCTGCGCT CTTTTGGACCGAGTTACCCAC	(GT) ₂₁	1
PL92	AF406748	TCTGGCAAAGTCTCTGTCTG GTATTGCTCAGGCTATCTT	(CA) ₁₉ CGAAAA (CA) ₂	1
PL115	AF406749	ACCTTCCTATGGATCGCTGG TTCAATGAAGTCCTCTCTGG	(CA) ₃₃	1
PL142	AF406750	ACTGTTACCTCGTCCTACGT GAGATGAAAAAGACTGTTGC	(GT) ₆ TT (GT) ₂₂	1

Tabla II.3. Continuación.

Locus	Nº de acceso	Secuencia del cebador (5'-3')	Motivo de repetición	Ref
PL167	AF406751	CAGACAAAATGAATGACAGG CAAGCAGCAGTCCCAGCA	(GT) ₂₄	1
Hhi1	AJ270779	GGAATAAAGAAGGGGGTGCG TGTGGTGGGTGGCAGTGGCTA	(GT) ₄ GA(GT) ₆ GA(GT) ₅ (GA) ₂ (GT) ₇ (GA) ₂	2
Hhi3	AJ270780	TCAGACAGGAAGGAAGTTTGG CCTCTCGGAATCACACACAG	(CA) ₃₂	2
Hhi51	AJ270781	TTGAGCCAGTTACAGAGAAGC ACTGTATCCTCTGTTTACATCCA	(TG) ₈ AG(TG) ₅	2
Hhi53	AJ270783	ACCAACAGTGACACATAGATCCT ATGCTAATGGGCTCTAAAATC	(CA) ₂₉	2
Hhi55	AJ270784	CTTTTTCCTGAGACGCTTG TAACCGTTCCTCCACTGC	(GT) ₇ TT(GT) ₈ (GA) ₄	2
Hhi56	AJ270785	CACCAAAGACAGATGAAGCA CTACACTATCAGCAGCCCAG	(GT) ₂ AT (GT) ₁₂	2
Hhi57	AJ270786	GATTGCTGCTGTTGCCTC TCCGCTGCTCCCTCTA	(CA) ₂ CT(CA) ₂ CT (CA) ₃ CT(CA) ₅ CTCAGC (CA) ₅ GA(CA) ₅ GA(CA) ₁₁ GA (CA) ₅ GA(CA) ₄	2
Hhi59	AJ270787	GAGTGAGAGAAACCAAAAGGC GCGAGGGAAGAGAGGAACAAC	(CT) ₂ (GT) ₁₂	2
Hhi60	AJ270788	CAGACAAAACTCACACACGCTC GCAAGTTTCATATAGGGGGTCAGAC	(CA) ₁₂	2
Hhi63	AJ270789	TCTCTATGTTTGCTGCCACCTTC TCGACCATCGTTTGAATCTTTTG	(CA) ₂₈ G(CA) ₂ (CGCA) ₉	2

Las condiciones de amplificación de los loci microsatélite se pusieron a punto en un termociclador de gradiente térmico (MJ Research). Este aparato permite testar un gradiente de temperaturas de anillamiento (T_a) a lo largo de cada fila en la placa de PCR, de tal forma que en cada pocillo de dicha placa se puede ensayar una temperatura de anillamiento distinta. En este trabajo se ensayó un rango de $\pm 4^\circ\text{C}$ respecto a la T_a propuesta para el cebador en la publicación original. Además, para cada T_a se probaron distintas concentraciones de magnesio.

Los cebadores que mostraron productos de amplificación con posible perfil microsatélite fueron posteriormente amplificados en una muestra de cinco rodaballos adultos de la población Cantábrica y analizados en gel de agarosa. Dos parejas de cebadores, la PL06 y la PL167, fueron marcadas con fluorescencia (5-N-N-ditol-tetrametilindodicarbocianina, Cy5) y sus productos de

amplificación se analizaron en el secuenciador automático ALF-express (*Automatic Laser Fluorescent DNA Sequencer*, Amersham Pharmacia Biotech; para más detalles consultar apartado II.2.1.3.1).

II.2.1.1.2. Librerías enriquecidas

A partir del ADN de un único espécimen de rodaballo se construyó una librería genómica parcialmente enriquecida en secuencias microsatélite. Dicha librería se obtuvo en la Universidad de Cambridge siguiendo el protocolo FIASCO (*Fast Isolation by AFLP of Sequences Containing repeats*, Zane *et al.*, 2002) con una sonda dinucleotídica (AC)₁₇. Los productos amplificados por PCR con un tamaño en torno a los 200 pares de bases (pb) fueron ligados en vectores pGEM-T (Promega) y transformados en células ultracompetentes XL2-Blue (Stratagene). En total, se picaron 84 colonias recombinantes y se sembraron en placas microtiter, donde se hicieron crecer con 100 µl de medio 2xTY y ampicilina a 37 °C durante toda la noche. Para su análisis, los insertos fueron amplificados por PCR usando cebadores de los vectores de clonación T7 y SP6 a fin de chequear sus tamaños. Setenta y ocho productos de la PCR se purificaron y secuenciaron siguiendo el protocolo del kit de secuenciación ABI Prism BigDye™ Terminator v3.1 Cycle. La secuenciación se llevó a cabo en los secuenciadores ABI 377 y ABI 3100 (Applied Biosystems).

II.2.1.2. Puesta a punto de las condiciones de amplificación PCR.

Para cada supuesto locus microsatélite, se diseñaron los correspondientes cebadores usando el programa PRIMER 3.0 (Rozen y Skaletsky, 2000). Los parámetros para el diseño de los cebadores se orientaron a la obtención de productos de PCR en torno a los 100-125 pb a partir de cebadores con un tamaño de 20 pb y una temperatura de anillamiento entre 55 y 62°C. Cada locus se ensayó en dos especímenes adultos de rodaballo procedentes de una población natural del Cantábrico. El protocolo de amplificación por PCR se detalla en el Apéndice IV.

La T_a se ensayó a lo largo de cada columna en la placa de PCR y la concentración de Mg^{+2} en cada fila. Las amplificaciones resultantes se visualizaron por electroforesis en un gel de agarosa al 2% en 35 ml de TBE 1x (0,1 M Tris-borato; 0,002 M EDTA, pH 8,0) y bromuro de etidio al 7%. Los carriles en los extremos del gel se cargaron con el marcador de tamaño 100 pb.

Para cada microsatélite se seleccionaron las condiciones de amplificación más astringentes que reprodujeran un perfil limpio de bandas espúreas. En aquellos loci donde la amplificación resultó poco específica se reajustaron las condiciones, bien disminuyendo la concentración de $MgCl_2$ o bien variando el rango de temperaturas de anillamiento.

II.2.1.3. Identificación de loci polimórficos.

El polimorfismo de los nuevos microsatélite se determinó mediante su amplificación en muestras de ADN de 22 individuos adultos de rodaballo, procedentes de una población natural del Cantábrico. La mezcla de reacción para la PCR se preparó siguiendo el protocolo descrito anteriormente, con la diferencia de que uno de los cebadores de la pareja se marcó con Cy5. La amplificación se llevó a cabo en un termociclador PTC-200 (MJ Research) aplicando el programa descrito en el apéndice IV, si bien, para cada pareja de cebadores se aplicó su temperatura de anillamiento específica. Las amplificaciones se comprobaron en un gel de agarosa al 2% y tinción con bromuro de etidio.

La caracterización de tamaño de los fragmentos resultantes de la amplificación de los loci microsatélite se llevó a cabo en el secuenciador automático ALF-express (Amersham Pharmacia Biotech).

II.2.1.3.1. Detección automatizada de fragmentos de ADN.

El secuenciador automático ALF-express (Amersham Pharmacia Biotech) permite la caracterización de tamaño de fragmentos de ADN mediante

una electroforesis vertical de alta resolución en un gel de poliacrilamida (ReproGel™ High Resolution Solution A, Amersham Pharmacia Biotech). El sistema incluye una fuente de electroforesis, un detector de fluorescencia por láser y un software informático para el procesamiento de los datos.

En las 40 posiciones de carga del gel de poliacrilamida se introduce una mezcla de reacción que incluye: dos marcadores internos de tamaño que flanquean al loci microsatélite; el ADN producto de la amplificación PCR; y tampón azul de dextrano. Los 2 pocillos de los extremos y al menos uno de los situados en el interior del gel, se cargan con marcadores externos, es decir, series de estándares de tamaños conocidos que se usan como referencias externas para inferir la longitud de los alelos microsatélite.

Cuando en la migración electroforética los productos de la PCR atraviesan el láser, los cebadores marcados emiten una señal fluorescente que es detectada por un sensor, que transforma la señal en un perfil de picos visibles en la pantalla. La primera señal en ser detectada corresponde a los fragmentos más pequeños formados por la asociación de cebadores o dímeros, a continuación en orden de tiempo se van detectando los marcadores de menor tamaño, los alelos de los loci microsatélite y finalmente, los marcadores internos de mayor tamaño. Estos perfiles de curvas se repiten para cada una de las muestras (Figura II.2).

El tamaño de los alelos microsatélite es estimado en pb automáticamente por el programa ALFwin Fragment Analyser 1.01 (Amersham Bioscience) en base al tiempo de migración y al tamaño conocido de los marcadores internos y externos.

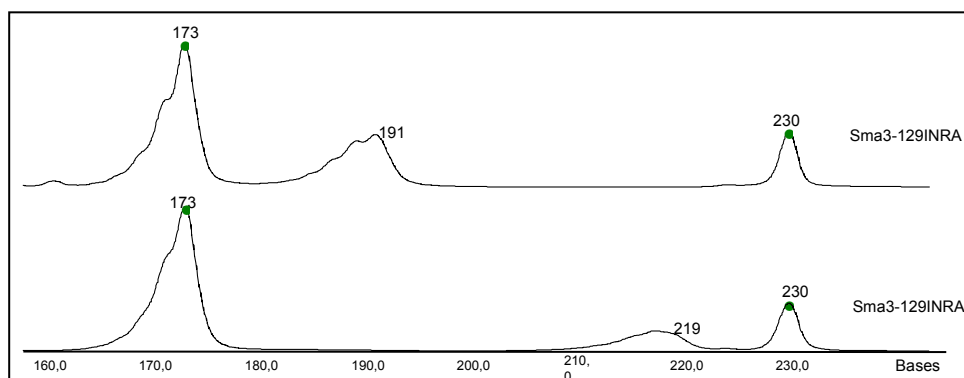


Figura II.2. Patrones de electroforesis vertical para el microsatélite Sma3-INRA en el ALF-Express (Amersham Pharmacia Biotech). Se muestran las carreras de dos individuos diferentes, los picos en los extremos corresponden a los marcadores de tamaño externos (de 173 pb y 230 pb, respectivamente). Los picos internos corresponden a los alelos de cada individuo, el primero homocigótico para el alelo 191 pb y el segundo, homocigótico para el alelo 219 pb.

II.2.1.4. Evaluación de la diversidad genética.

La diversidad genética de los loci polimórficos se evaluó a partir de los datos genotípicos de los 22 individuos de la muestra cantábrica. Las frecuencias alélicas y los niveles de diversidad genética (número de alelos por locus y heterocigosis esperada; Nei 1987) se calcularon utilizando el programa GENEPOP 3.1 (Raymond y Rousset, 1995). El test de ajuste al equilibrio Hardy-Weinberg (HWE) se realizó usando el método de cadenas de Markov implementado en el mismo paquete estadístico. El desequilibrio de ligamiento genotípico entre todos los pares de loci se testó mediante un test exacto usando el mismo programa. Posteriormente, se aplicó la corrección secuencial de Bonferroni para múltiples test (Rice, 1989).

II.2.1.5. Genotipado de marcadores microsatélite en la familia de referencia

Cada locus microsatélite (31 procedentes de la literatura y 5 obtenidos en el presente estudio) se amplificó a partir de 50 ng del ADN genómico de la hembra H8. Las amplificaciones se llevaron a cabo según el protocolo PCR descrito en el apéndice IV, si bien las concentraciones de $MgCl_2$ y las temperaturas de anillamiento se modificaron de acuerdo con los requerimientos

de cada locus. La primera evaluación de los productos de la PCR se realizó mediante electroforesis en gel de agarosa al 2% a fin de comprobar el éxito de la amplificación y de estimar la cantidad de ADN amplificado a cargar en el secuenciador automático para el análisis preciso de fragmentos de ADN.

Los loci heterocigóticos en la hembra H8 fueron seleccionados como marcadores para la construcción del mapa genético. Estos loci fueron amplificados en los 50 embriones haploides de la familia H8. En cada reacción de PCR, se utilizaron 50 ng del ADN previamente amplificado con el kit de amplificación genómica GenomiPhi™.

II.2.2. Tecnología AFLP

La tecnología AFLP implica tres pasos: (i) digestión enzimática de ADN genómico y ligamiento de adaptadores oligonucleotídicos; (ii) amplificación selectiva de un subgrupo de los fragmentos resultantes de la restricción; y (iii) análisis genotípico de los fragmentos amplificados.

II.2.2.1. Optimización de la metodología AFLP

La metodología AFLP se aplicó según el protocolo original de Vos *et al.* (1995) con las modificaciones propuestas por Papa *et al.* (2005). El protocolo de Vos y colaboradores (1995) se diseñó para el uso de cebadores marcados con radioactividad y para la resolución de las bandas en geles de secuenciación. Papa *et al.* (2005) consideraron la identificación de marcadores AFLP mediante marcaje fluorescente y electroforesis capilar, y propusieron las siguientes modificaciones respecto al protocolo original:

i) Enzimas de restricción. Frente a la combinación clásica de endonucleasas *EcoRI/MseI* ampliamente utilizada en vegetales (Vos *et al.*, 1995), Papa y colaboradores (2005) proponen el uso de la combinación *EcoRI/TaqI* para favorecer la detección de polimorfismos en los patrones AFLP derivados de genomas animales. Considerando la diana de corte de *TaqI*

(TCGA) y la mayor propensión de los dinucleótidos CpG a la transición de citosinas metiladas a timinas, el uso de esta enzima podría revelar un mayor número de polimorfismos derivados de esta mutación (Vos y Kuiper, 1996; Ajmone-Marsan *et al.*, 1997).

ii) Tiempos de digestión y ligamiento. Para reducir el ruido de fondo, Papa *et al.* (2005) proponen incrementar los tiempos de digestión del ADN y de ligamiento de los adaptadores respecto al protocolo original propuesto por Vos *et al.* (1995). Además, las reacciones de restricción se llevan a cabo a diferentes temperaturas para las enzimas *TaqI* (65°C) y *EcoRI* (37°C), a fin de evitar la inactivación por calor de ésta última.

II.2.2.2. Protocolo AFLP

ADN genómico. A fin de evitar posibles errores de genotipado derivados de un ADN genómico fragmentado, para el desarrollo del protocolo AFLP se utilizó el ADN original de alto peso molecular, extraído a partir de los embriones haploides, mientras que las amplificaciones de este ADN obtenidas con el protocolo GenomiPhi™ se reservaron para el genotipado de los marcadores microsatélite, menos exigentes respecto a la calidad del ADN de partida.

En cuanto a la cantidad de ADN inicial para la aplicación del protocolo AFLP, no parece existir una medida establecida. Vos *et al.* (1995) recomendaban utilizar entre 200 y 400 ng de ADN genómico para la restricción. Papa *et al.* (2005) proponen incrementar la cantidad inicial de ADN a 600-700 ng, especialmente si se trabaja con ADN parcialmente degradado. Sin embargo, otros autores describen la obtención exitosa de perfiles AFLP a partir de cantidades de ADN genómico no superiores a 50 ng (Muller y Wolfenbarger, 1999; Liu *et al.*, 2003). En todo caso, la disponibilidad del ADN inicial se ve limitada por la naturaleza de las muestras. Cuando se trabaja con embriones haploides, la cantidad total de ADN genómico que se puede extraer es reducida. Teniendo en cuenta esta limitación y la precaución de reservar ADN

de cada espécimen, se ensayó la obtención de perfiles AFLP a partir de cantidades iniciales de 100 y 200 ng de ADN genómico. La comparación de los perfiles AFLP obtenidos en cinco embriones haploides para una y otra cantidad de ADN no reveló diferencias significativas y finalmente, se decidió partir de la cantidad menor (100 ng) para el desarrollo del protocolo AFLP en los restantes miembros de la familia H8.

Restricción-Ligamiento y Amplificación AFLP. Los protocolos de corte con los enzimas de restricción y el posterior ligamiento de los adaptadores, junto con la amplificación de los AFLP se detallan en el apéndice V.

II.2.2.3. Selección de cebadores AFLP

El diseño de los cebadores AFLP dependerá principalmente de la secuencia nucleotídica de los adaptadores, además de las dianas de corte y de las bases selectivas. Las secuencias de los adaptadores y de los cebadores usados en este trabajo reproducen los diseños propuestos por Vos *et al* (1995), con las modificaciones requeridas dependiendo de las secuencias de los sitios de restricción (Tabla II.4).

Para la amplificación selectiva, 22 combinaciones de estos cebadores se testaron en un panel de 20 embriones haploides de la familia H8. Las combinaciones analizadas se señalan con un punto en la Figura II.3. Sólo las combinaciones de cebadores que permitieron detectar al menos tres polimorfismos fueron seleccionadas para el genotipado de los restantes miembros de la familia H8.

Tabla II.4. Secuencias de los adaptadores y cebadores usados para el análisis AFLP. La nomenclatura de los cebadores corresponde a la casa comercial *Keygene* (ver apartado II. 2.2.6). En las secuencias de los mismos, se indican en **negrita** las bases selectivas.

	Código (según <i>Keygene</i>)	Secuencia
Adaptadores <i>EcoRI</i>	<i>Eco</i> cadena superior	5' CTCGTAGACTGCGTACC 3'
	<i>Eco</i> cadena inferior	3' CATCTGACGCATGGTTAA 5'
Adaptadores <i>TaqI</i>	<i>Taq</i> cadena superior	5' GACGATGAGTCCTGAC 3'
	<i>Taq</i> cadena inferior	3' TACTCAGGACTGGC 5'
Cebador <i>EcoRI</i>	E01	5' GACTGCGTACCAATT CA 3'
	E32	*5' GACTGCGTACCAATT CAAC 3'
	E33	*5' GACTGCGTACCAATT CAAG 3'
	E40	*5' GACTGCGTACCAATT CAGC 3'
	E42	*5' GACTGCGTACCAATT CAGT 3'
Cebador <i>TaqI</i>	T01	5' GATGAGTCCTGACCG AA 3'
	T02	5' GATGAGTCCTGACCG AC 3'
	T32	5' GATGAGTCCTGACCG AAAC 3'
	T33	5' GATGAGTCCTGACCG AAAG 3'
	T37	5' GATGAGTCCTGACCG AAACG 3'
	T38	5' GATGAGTCCTGACCG AACT 3'
	T49	5' GATGAGTCCTGACCG ACAG 3'
	T50	5' GATGAGTCCTGACCG ACAT 3'
	T54	5' GATGAGTCCTGACCG ACCT 3'
	T58	5' GATGAGTCCTGACCG ACGT 3'

Los cebadores marcados con fluorescencia se indican por un asterisco que precede al extremo 5

		Cebador <i>TaqI</i>							
		T32	T33	T37	T38	T49	T50	T54	T58
Cebador <i>EcoRI</i>	E32			•		•	•	•	
	E33		•	•	•	•		•	•
	E40	•	•	•		•	•	•	
	E42	•	•	•		•	•		•

Figura II.3. Combinaciones de cebadores selectivos para el análisis AFLP. Se indican con un punto las combinaciones de cebadores analizadas en la hembra H8 y una submuestra de su descendencia haploide (20 embriones). Sin señalar aparecen las combinaciones no testadas.

II.2.2.4. Detección automatizada de AFLPs

La detección de AFLPs se llevó a cabo por electroforesis capilar en el secuenciador automático CEQ 2000™ de Beckman-Coulter (Figura II.4). El secuenciador está equipado con una serie de ocho capilares de separación de 33 mm cada uno y un sistema láser de detección de fragmentos. El sistema de detección actúa en un modo similar al del secuenciador automático ALF-express, sin embargo, la electroforesis capilar es más sensible a pequeñas

cantidades de ADN y puede diferenciar con precisión fragmentos de ADN que se distinguen en no más de una base de longitud.

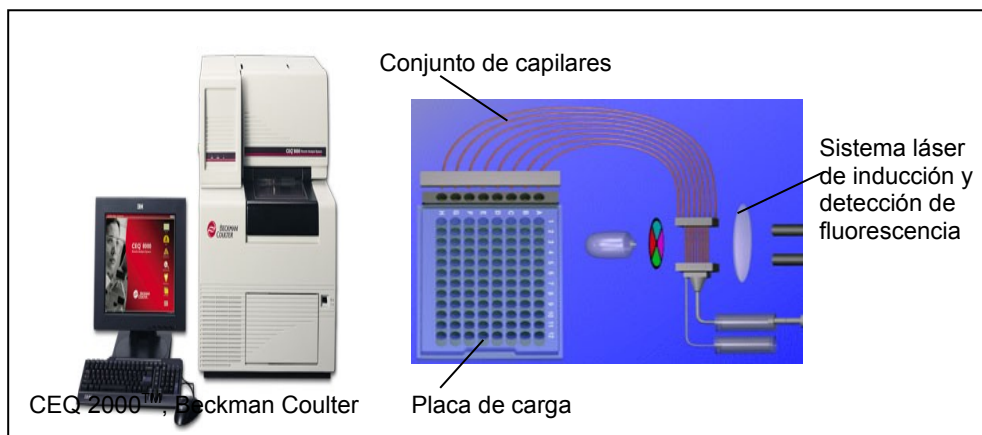


Figura II.4. Esquema básico del funcionamiento del CEQ 2000. El ADN cargado en la placa es automáticamente inyectado en los capilares rellenos de poliacrilamida. Cuando las muestras desnaturalizadas atraviesan el láser, el fluoróforo del cebador marcado se excita, emitiendo una señal reconocible por el sistema.

Los fragmentos detectados se muestran como un perfil de picos. El análisis de fragmentos de ADN es llevado a cabo mediante la aplicación "Fragment analysis module" del secuenciador. Este módulo procesa los datos de la plataforma CEQ 2000 y proporciona los tamaños y la información alélica para los picos detectados. Los tamaños en pares de bases de los fragmentos asociados con los picos se estiman a partir de una curva de calibrado, la cual se construye enfrentando la información conocida del tamaño de los fragmentos del estándar y del tiempo de migración de los mismos.

II.2.2.5. Análisis de los perfiles AFLP. Software GENOGRAPHER

Para facilitar el análisis de los perfiles AFLP e identificar los potenciales polimorfismos, los datos cromatográficos del CEQ 2000 se transformaron en un patrón de bandas de un gel virtual. Para ello se usó el software GENOGRAPHER (Ver. 1.6.0, Benham J.J. Montana State University 2001, Figura II.5). Los archivos de datos del CEQ2000 se importan a GENOGRAPHER, que considera en cada perfil AFLP la información relacionada con la altura de pico, resolución y

movilidad, y transforma la señal de cada fragmento en una banda de un gel virtual. Los polimorfismos AFLP se detectan entonces como presencia/ausencia de banda.

Usando el software GENOGRAPHER, se estudiaron los fragmentos presentes en el gel en un rango de 70 a 550 pb. Este intervalo excluye las bandas de tamaños extremos a fin de evitar errores de genotipado derivados de la técnica. Los fragmentos más grandes normalmente son escasos y amplifican con mayor dificultad que los fragmentos pequeños, además la posibilidad de detectarlos dependerá en gran medida de la calidad del ADN de partida (Qi y Lindhout, 1997). Todo ello se traduce en un defecto de reproducibilidad entre los individuos que podría interpretarse como falsos polimorfismos. Por otro lado, los fragmentos más pequeños, en torno a las 50 pb o menores, a pesar de caracterizarse por un alto valor de fluorescencia, a menudo resultan desalineados por cuestiones relacionadas con el funcionamiento del aparato de secuenciación. Este desfase en el alineamiento puede llevar a interpretaciones incorrectas en cuanto a la presencia o no de polimorfismos. A fin de evitar estas fuentes de error, sólo las bandas comprendidas en el rango establecido (70 a 550 pb) se consideraron informativas. Dentro de este rango, se excluyeron del análisis las bandas monomórficas, las no reproducibles o las de menor resolución debido a la presencia de rastro de fondo (*smears*) o a una migración demasiado próxima a la(s) banda(s) vecina(s). Para considerar los fragmentos más pequeños de entre 70 y 200 pb, se estableció 2000 cnts como límite de intensidad de fluorescencia. En picos de mayor tamaño, a partir de 300 pb, se consideraron también intensidades menores, especialmente en los fragmentos de alto peso molecular.

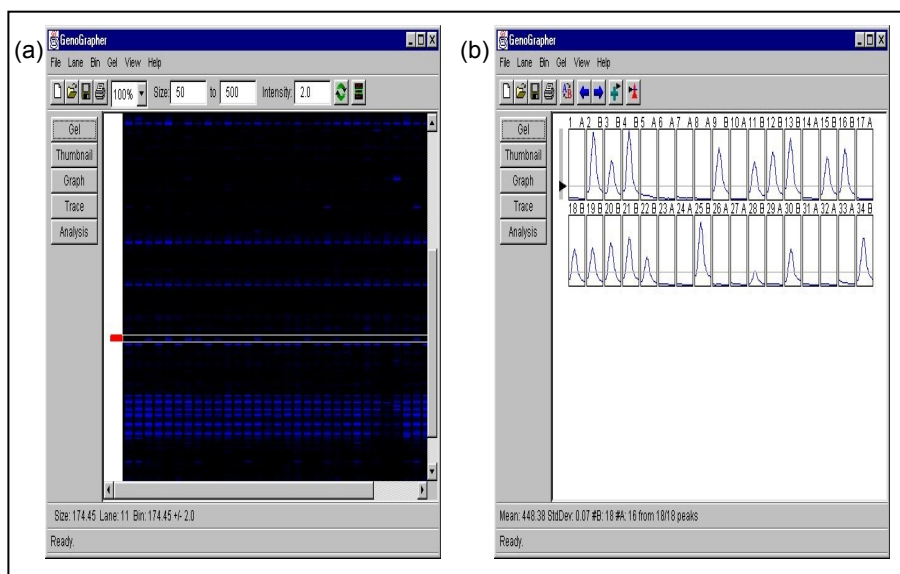


Figura II.5. Dos opciones gráficas de Genographer: (a) representa un patrón de bandas AFLP, mientras que (b) muestra el perfil de los picos correspondientes a cada individuo para una banda de un tamaño determinado. Estos picos reproducen los detectados por el CEQ 2000.

Finalmente, GENOGRAPHER da la opción de convertir estos geles en un formato de texto estándar. Para cada marcador, la presencia/ausencia de banda se codifica como genotipos A/H, respectivamente, siendo éste el formato reconocible por MAPMAKER (Lander *et al.*, 1987; para más detalles consultar apartado II.3.2.1). De este modo, los fenotipos codificados por GENOGRAPHER pueden importarse directamente al programa de mapeo.

II.2.2.6 Nomenclatura AFLP

No existe una norma establecida acerca de la nomenclatura de los marcadores AFLP. Sin embargo, la mayoría de los trabajos de caracterización de AFLPs nombran estos marcadores de acuerdo a la combinación de cebadores y al tamaño de los fragmentos resultantes (Kocher *et al.*, 1998; Cervera *et al.*, 2001; Liu *et al.*, 2003). Siguiendo este criterio, la denominación de cada marcador AFLP identificado en la familia H8 incluye la combinación de cebadores utilizada (Tabla II.4), seguida de un número que se refiere al tamaño en pares de bases del fragmento identificado. Así por ejemplo, un

marcador de 250 pb identificado a partir de la combinación de cebadores E32/T37 se denominará E32/T37.250

II.2.3. Test de conformidad con la segregación mendeliana.

Antes de iniciar el análisis de ligamiento, se comprobó en todos los marcadores el ajuste a las proporciones esperadas de segregación mendeliana. Para ello se calcularon los valores de chi-cuadrado (χ^2 , g.l = 1) de la segregación alélica de todos los marcadores microsatélite y AFLP genotipados en la descendencia. Estos valores se testaron respecto a las proporciones esperadas en una segregación mendeliana 1:1. Los marcadores que se desviaron significativamente ($P < 0,01$) de las proporciones esperadas se reexaminaron en busca de posibles errores de genotipado y fueron eliminados aquellos donde la desviación persistió tras la corrección. Los marcadores que mostraron distorsión de segregación $0,01 < P < 0,05$ fueron descartados para el desarrollo inicial del mapa y sólo se ordenaron en los grupos de ligamiento, una vez establecida la posición de los marcadores acordes con la segregación mendeliana.

II.2.4. Fuentes de error de genotipado.

Trabajos basados en la reproducibilidad de los experimentos de caracterización de marcadores microsatélite y AFLP han estimado tasas de error de genotipado del 0,8% y 2,6% para unos y otros marcadores, respectivamente (Mueller y Wolgenbarger, 1999; Bonin *et al.*, 2004). Las causas de error en estos procesos son múltiples y no siempre se pueden determinar con exactitud, sin embargo, ciertas características intrínsecas a la metodología AFLP hacen a este procedimiento más susceptible a la introducción de errores. Bonin *et al.* (2004) diferencian en la tasa de error las disparidades ocasionadas por el procedimiento técnico y las introducidas por factores humanos. Entre las causas que afectan al procedimiento técnico se podrían incluir:

- La calidad y concentración del ADN de partida. Ambos aspectos van a influir en gran medida en las etapas iniciales de digestión y ligamiento del protocolo AFLP. La digestión de un ADN degradado o fracturado generará fragmentos adicionales, generalmente de alto peso molecular, que alterarán el patrón de picos o bandas AFLP, pudiendo ocasionar falsos polimorfismos (Vos *et al.*, 1995; Mueller y Wolgenbarger, 1999). Por otro lado, cuando la cantidad de ADN de partida es particularmente escasa (en torno a 1pg, datos de Vos *et al.*, 1995) los fragmentos originados por la restricción enzimática no se distribuirán al azar durante el proceso de amplificación, dando lugar a diferencias significativas en la intensidad de las bandas amplificadas.
- Defectos de migración. Entre los factores técnicos que influyen más habitualmente en la tasa de error, destacan ligeras desviaciones en la migración electroforética de fragmentos homólogos (Bonin *et al.*, 2004). Estas desviaciones pueden inducir a la puntuación de falsos polimorfismos ocasionados por la distinta movilidad de las bandas en carriles contiguos.
- Presencia de bandas “parásitas” prodecentes de la digestión del ADN de microorganismos que hayan podido contaminar las muestras.
- Homoplasia. Esta fuente de imprecisión en el genotipado de marcadores AFLP raramente es tomada en cuenta, sin embargo trabajos recientes señalan la comigración de fragmentos no homólogos como causa potencial de error (O’Hanlon y Peakall, 2000; Vekemans *et al.*, 2002; Bonin *et al.*, 2004). En un trabajo previo, Nikaido *et al.* (1999) previenen sobre el efecto de la presencia de bandas complejas en la distorsión de la segregación mendeliana que exhiben algunos loci AFLP.

- Amplificaciones inespecíficas. Los productos de este tipo de amplificaciones originan rastros de fondo que en ocasiones pueden dificultar la correcta interpretación de presencia o ausencia de banda en un patrón AFLP. Este efecto se puede corregir haciendo más restrictivas las condiciones de amplificación y siendo cuidadosos en el diseño de los cebadores.

Todas las variables comentadas pueden afectar al patrón de bandas AFLP, llevando a introducir errores en el genotipado de las muestras. Así, la experiencia del investigador responsable del análisis es fundamental en la correcta identificación de un genotipo. Bonin *et al.* (2004) señalan la subjetividad en la interpretación de las diferencias técnicas como la principal fuente de error en el proceso de genotipado de marcadores AFLP. Son especialmente dependientes de la subjetividad del investigador la interpretación de la disparidad en la intensidad de los picos como presencia o ausencia de banda y la lectura de las desviaciones de migración.

II.2.4.1. Precauciones metodológicas en la caracterización de marcadores AFLP

En este trabajo, a fin de minimizar la tasa de error derivada de variaciones técnicas, se tomaron una serie de precauciones al aplicar el protocolo AFLP. La calidad del ADN obtenido a partir de la extracción con fenol: cloroformo: alcohol isoamílico se comprobó mediante electroforesis en gel de agarosa al 1%. Aquellas muestras en las que se observó un ADN de baja calidad fueron descartadas para posteriores análisis.

El ADN de alto peso molecular de los 20 individuos seleccionados para el ensayo de los cebadores fue sometido al protocolo AFLP. Tras la amplificación selectiva, la reproducibilidad del patrón de bandas AFLP originado por cada una de las combinaciones de cebadores fue comprobada entre los 20 individuos amplificados. Una vez seleccionadas las combinaciones de cebadores en función de los polimorfismos originados así como de la

fiabilidad de los perfiles AFLP, se completó el genotipado de la familia H8. Aquellos individuos en los que se apreciaron perfiles raros o desalineados fueron reanalizados y en los casos donde las anomalías no se corrigieron tras la repetición, los individuos fueron descartados. Así, se valoraron los descendientes de la hembra H8 hasta seleccionar una muestra de 50 individuos.

Además de estas precauciones, en el caso del mapeo genético, la comprobación estadística del ajuste de la segregación alélica a las proporciones mendelianas contribuye también a identificar marcadores con un genotipado inconsistente.

II.2.4.2. Dos fuentes particulares de errores de genotipado: amplificaciones inespecíficas y presencia de bandas complejas.

Como se ha comentado anteriormente existen dos tipos de diferencias técnicas que pueden llevar a errores de genotipado y que inciden particularmente en la distorsión de la segregación mendeliana: la amplificación de una misma banda por más de una pareja de cebadores y la asunción no verificada de que los fragmentos de igual tamaño son realmente homólogos.

II.2.4.2.1. Evaluación de posibles amplificaciones inespecíficas

A fin de evaluar la presencia de posibles amplificaciones inespecíficas entre los productos PCR de una determinada combinación de cebadores, se estudió la coincidencia de tamaño entre los fragmentos amplificados por combinaciones de cebadores similares en cuanto a su composición de bases nucleotídicas.

Considerando que cuanto más parecidas son las secuencias nucleotídicas de los cebadores, mayor es la probabilidad de que ocurra un error de emparejamiento (*mismatch*), se comparó el porcentaje medio de polimorfismos comunes resultante de combinaciones de cebadores parecidas y el resultante de combinaciones de cebadores más diferenciados en cuanto a

su composición de bases. Se consideraron combinaciones de cebadores parecidas aquellas que difieren en su composición en sólo una base, por ejemplo, E32/T37 y E33/T37, entre los cuales coinciden cinco de las seis bases selectivas incluidas en el extremo 3': AAC/ACG vs. AAG/ACG.

Si durante el anillamiento de los cebadores ocurren emparejamientos inespecíficos, es de esperar que el porcentaje de bandas coincidentes en tamaño sea mayor entre las parejas de cebadores de composición nucleotídica parecida. Si por el contrario, la coincidencia en tamaño se debe al azar, el porcentaje de bandas comunes entre cebadores parecidos y bien diferenciados no debería ser significativamente distinto. Siguiendo este razonamiento, los tamaños de los productos amplificados por cada una de las 17 combinaciones de cebadores AFLP se compararon, uno a uno, con los productos amplificados por las 16 combinaciones restantes.

II.2.4.2.2. *Caracterización de bandas complejas. Homoplasia*

La presencia de bandas complejas entre los perfiles AFLP se testó aplicando el procedimiento descrito por O'Hanlon y Peakall (2000). Dicho método se basa en la adición de bases selectivas (T+4) al extremo 3' de uno de los cebadores AFLP de una pareja determinada (E+3/T+3). Puesto que un sitio dado de ADN puede tener una de las 4 bases (A, T, G o C), añadiendo una base conocida a uno de los cebadores AFLP se reducirá 4 veces el número de fragmentos amplificados (Liu y Cordes, 2004), de tal forma que los perfiles resultantes serán un subconjunto del patrón de bandas original. La amplificación de un fragmento de un determinado tamaño por más de un cebador T+4 demostraría la presencia de bandas complejas, constituidas por fragmentos de igual tamaño procedentes de diferentes regiones del genoma (homoplasias) (Nikaido *et al.*, 1999).

En este trabajo, el patrón de bandas originado por una pareja de cebadores con seis bases selectivas (E33/T49) se comparó con el resultante

de 4 combinaciones de cebadores con 7 bases selectivas (E33/T49+N). La composición nucleotídica de cada cebador se muestra en la Tabla II.5.

Tabla II.5. Cebadores AFLP utilizados en la evaluación de homoplasia de tamaño. Se indica la nomenclatura y secuencia de bases de cada cebador.

Cebador	Código	Secuencia
<i>EcoRI</i>	E33	*5' GACTGCGTACCAATTC AAG 3'
<i>TaqI</i>	T49	5' GATGAGTCCTGACCGA CAG 3'
	T49 + A	5' GATGAGTCCTGACCGA CAGA 3'
	T49 + T	5' GATGAGTCCTGACCGA CAGT 3'
	T49 + C	5' GATGAGTCCTGACCGA CAGC 3'
	T49 + G	5' GATGAGTCCTGACCGA CAGG 3'

II.3. CONSTRUCCIÓN DEL MAPA GENÉTICO DE LIGAMIENTO

Una vez seleccionada la familia de referencia y caracterizados los marcadores moleculares, la construcción del mapa genético implica las siguientes etapas: (i) detección de ligamiento entre marcadores; (ii) ordenación de los marcadores ligados; y (iii) estimación de distancias entre los mismos. En un mapa de moderada densidad, el volumen de datos que debe manejarse en estos cálculos hace prohibitiva la identificación de ligamiento a partir del cómputo de recombinantes, por lo que se ha desarrollado un gran número de programas específicamente diseñados para el análisis computacional de la asociación entre marcadores y su posterior ordenación dentro de los grupos de ligamiento.

II.3.1. Softwares de análisis de ligamiento

La mayoría de los programas que se utilizan en la construcción de mapas genéticos están disponibles en la página web del Rockefeller Institute, <http://linkage.rockefeller.edu/soft>. Básicamente, estos softwares usan los mismos estadísticos básicos, llamados de análisis de dos puntos para estimar tasas de recombinación y testar ligamiento entre pares de marcadores. Sin embargo, los programas difieren en varios conceptos, como el tipo de población de mapeo que pueden analizar (retrocruzamiento, F_2 , pedigris clásicos, etc.); los algoritmos utilizados para estimar las frecuencias de

recombinación multipunto; el método empleado para ordenar los marcadores; y en algunas características informáticas (formato de los datos, *interface*, características gráficas, etc.). En la Tabla II.6 se citan algunos de los programas más populares de mapeo genético, incluyendo una pequeña descripción de cada uno.

Tabla II.6. Programas de análisis de ligamiento. Breve descripción y referencia.

Nombre de la herramienta	Descripción	Referencia
LIPED	Herramienta de análisis de dos puntos. Cálculo de puntuación LOD. Fue el cebador programa disponible para computar numéricamente verosimilitudes en pedigrís grandes.	Ott, 1974; 1976
LINKAGE	Paquete de herramientas que incluye análisis multipunto. Cálculo de puntuación LOD. Existen dos versiones: un conjunto de programas para pedigrís generales y una versión especializada para el análisis de marcadores codominantes en el panel CEPH de familias de referencia.	Lathrop y Lalouel, 1984; Dausset <i>et al.</i> , 1990
CRI-MAP	Desarrollado inicialmente para analizar sistemas de marcadores codominantes en familias pequeñas, se ha extendido a pedigrís más generales y a sistemas de dominancia simple.	Goldgar <i>et al.</i> , 1989
GENEHUNTER	Herramienta que incluye análisis de ligamiento paramétricos y no paramétricos en un modo unificado.	Kruglyak <i>et al.</i> , 1996
MAPMAKER	Herramienta de análisis multipunto. Cálculo de puntuación LOD. Puede analizar datos con muchos loci pero en pedigrís pequeños y en cruces entre líneas controladas.	Lander <i>et al.</i> , 1987
JOINMAP	Es el programa más usado en proyectos de mapeo de especies vegetales. Puede trabajar prácticamente con todo tipo de poblaciones. La ordenación de marcadores se obtiene calculando frecuencias de recombinación por pares y las distancias de mapa se estiman por un procedimiento de mínimos cuadrados.	Stam, 1993; Stam y van Ooijen, 1995
MAPMANAGER QTX	Este programa permite cartografiar loci mendelianos de retrocruzamiento, líneas consanguíneas recombinantes o cruzamientos entre líneas	Manly <i>et al.</i> , 2001.

Sin embargo, no existe un programa óptimo para todas las prestaciones. En publicaciones recientes en proyectos de mapeo genético

(Gosselin *et al.*, 2002; Somers *et al.*, 2004; Benjafield *et al.*, 2005) programas clásicos como MAPMAKER (Lander *et al.*, 1987) coexisten con los últimos desarrollados como JOINMAP (Stam, 1993; Stam y van Ooijen, 1995) o GENEHUNTER (Kruglyak *et al.* 1996), usando uno u otro programa según los requerimientos de los distintos pasos en el proceso de construcción del mapa. En especies de peces, con frecuencia se ha recurrido a los programas MAPMAKER y MAPMANAGER (Manly *et al.*, 2001), en muchas ocasiones usando el primero para corroborar los resultados del segundo (Young *et al.*, 1998; Gates *et al.*, 1999; Nichols *et al.*, 2003; Walter *et al.*, 2004). En el presente estudio, para la construcción del mapa genético de rodaballo, se eligió el programa MAPMAKER al ser uno de los principales programas utilizado en la construcción de la mayoría de los mapas genéticos disponibles en peces (Kocher *et al.*, 1998; Wilson *et al.*, 2002; Gilbey *et al.*, 2004; Watanabe *et al.*, 2004; Kai *et al.*, 2005). Además, MAPMAKER permite trabajar con datos haploides codificándolos como un retrocruzamiento con algunas modificaciones simples, tal como se detalla en el apartado II.3.2.1.

II.3.2. MAPMAKER/EXP 3.0

MAPMAKER es un paquete de análisis de ligamiento diseñado para ayudar a construir mapas de ligamiento de marcadores segregantes, tanto en cruces experimentales como en poblaciones naturales. El programa original (Lander *et al.*, 1987), diseñado para el análisis de pequeños pedigrís, fue mejorado posteriormente para permitir análisis de ligamiento con medios hermanos (Kruglyak *et al.*, 1995), análisis de caracteres recesivos con familias nucleares (Kruglyak *et al.*, 1995) y análisis multipunto con muchos marcadores en pedigrís de tamaño moderado (Kruglyak *et al.*, 1996). Particularmente, MAPMAKER/EXP es una versión del paquete original diseñado específicamente para analizar datos de cruces experimentales. MAPMAKER/EXP lleva a cabo la estimación simultánea de todas las frecuencias de recombinación a partir de los datos genotípicos de marcadores dominantes, recesivos y codominantes. En este trabajo se usó el software MAPMAKER/EXP 3.0, una última versión (1993)

que incluye algunas mejoras con respecto al MAPMAKER/EXP original. Algunas de las principales ventajas de este programa son las siguientes:

- Permite un seguimiento detallado de cada paso en la construcción del mapa.
- Permite trabajar con fenotipos de loci codominantes, los cuales tendrían una difícil interpretación con programas diseñados para analizar marcadores dominantes.
- Incluye una opción de detección de errores. Dado un orden de mapa permite calcular *a posteriori* la probabilidad de que cada genotipo individual sea correcto o erróneo.
- MAPMAKER/EXP y especialmente la versión 3.0 puede analizar datos derivados de progenies de varios tipos de cruzamientos, incluyendo: retrocruzamiento F_2 , cruzamiento entre líneas F_1 , cruzamiento entre líneas F_3 (por autofecundación) y líneas endogámicas recombinantes (RIL: *Recombinant Inbred Lines*, por autofecundación completa o parcial). Hay también otros tipos de cruces que pueden ser analizados por MAPMAKER/EXP cuando su modelo de segregación genética se ajusta al de alguno de los cruces contemplados por el programa. Por ejemplo, la F_1 en haploides, donde los datos siguen el modelo genético del retrocruzamiento.
- Los archivos de MAPMAKER son compatibles con el programa MAPMAKER/QL para el mapeo de QTLs.
- Es un programa de libre acceso.

Sin embargo, se deben mencionar también algunas desventajas del programa:

- Las características gráficas de MAPMAKER son muy limitadas y con frecuencia se requieren otros programas para representar los grupos de ligamiento.

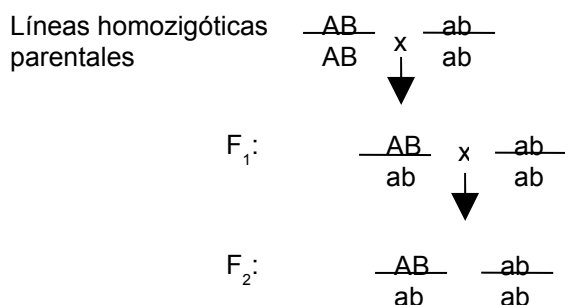
- Es un programa muy interactivo y requiere un alto conocimiento del proceso global de mapeo.

Debido a la naturaleza haploide del material biológico utilizado en este estudio, a lo largo de los próximos apartados se centrará la atención en los métodos numéricos para resolver retrocruzamientos con el software MAPMAKER/EXP 3.0.

II.3.2.1. Codificación para MAPMAKER de los datos genotípicos procedentes de familias ginogenéticas haploides.

Para poder identificar los eventos de recombinación se requiere conocer previamente la fase de ligamiento de los marcadores en los progenitores (Maliepaard *et al.*, 1997). Sin embargo, cuando se trabaja con cruces no controlados la fase de ligamiento de los marcadores en los progenitores es, *a priori*, desconocida. En estos casos, dicha fase debe deducirse a partir de la segregación en la propia descendencia.

Particularmente, en el caso de los embriones ginogenéticos haploides, sólo los alelos de uno de los progenitores (la madre) segregan en la descendencia. Algunos programas para análisis de ligamiento, como JOINMAP (Stam, 1993; Stam y van Ooijen 1995), pueden interpretar este tipo de datos, sin embargo, MAPMAKER (Lander *et al.*, 1987) no reconoce directamente este modelo genético. En este caso, la solución posible para el análisis de ligamiento es codificar la descendencia haploide como si se tratase de la F₂ de un retrocruzamiento en el que la madre sería el parental heterocigótico, mientras que el padre sería homocigoto recesivo para todos los loci. El siguiente esquema ilustra el modelo genético de retrocruzamiento:

Figura II.6. Retrocruzamiento F_2 .

En un retrocruzamiento, los alelos de la descendencia heterocigótica se encontrarán en disposición *cis* o de acoplamiento en los tipos parentales. Sin embargo, en una familia haploide originada a partir de una población natural, la configuración de los tipos parentales no necesariamente ha de ser la de acoplamiento, sino que pueden aparecer también en fase de repulsión o *trans*, situación que no se contempla en un retrocruzamiento. Para resolver esta cuestión, los genotipos de cada locus han de ser recíprocamente codificados, de tal forma que ambas fases, de acoplamiento o de repulsión puedan ser detectadas en el conjunto de datos analizados por MAPMAKER como un retrocruzamiento. Así por ejemplo, dado un marcador M con alelos m_1 y m_2 , heterocigoto en el progenitor, la segregación de cada uno de sus alelos en la descendencia haploide se codificará para el análisis en MAPMAKER como A (homocigótico en un supuesto retrocruzamiento) si portan el alelo m_1 y H (heterocigótico en el retrocruzamiento) si portan el alelo m_2 . Alternativamente, se introducirá en el archivo de datos un marcador Mr donde el alelo m_1 será codificado como H y el alelo m_2 como A. De esta forma, la existencia de ligamiento entre dos marcadores N y M será identificada por duplicado: en caso de que los alelos se presenten en configuración *trans* el ligamiento se detectará entre M y Nr y entre Mr y N , mientras que si los alelos se encuentran en configuración *cis*, se detectará ligamiento entre M y N y entre Mr y Nr . Así, para cada grupo de ligamiento el programa identificará un par de copias “especulares” igualmente informativas. A continuación se seleccionará aleatoriamente una de las copias para continuar con el proceso de construcción del mapa.

II.3.3. Identificación de los grupos de ligamiento. Análisis de dos puntos.

Un grupo de ligamiento se definen como un conjunto de marcadores donde cada marcador está ligado al menos a otro marcador del mismo conjunto. Como se adelantaba en el apartado de introducción, la identificación de los grupos de ligamiento se lleva a cabo mediante un análisis de dos puntos. A continuación se detallan las etapas en las que procede este análisis: estimación de las frecuencias de recombinación y test de ligamiento entre pares de marcadores.

II.3.3.1. Estimación de la frecuencia de recombinación. Método de Máxima Verosimilitud.

Si los tipos recombinantes y no recombinantes son directamente observables, una estima intuitiva de la fracción de recombinación puede expresarse como $\hat{r} = k/n$, siendo k el número de recombinantes observados y n el número de productos meióticos. Esta estima se conoce como *método directo* y se usó en los primeros intentos de análisis de ligamiento en humanos. Sin embargo, factores como la penetrancia incompleta o fases de ligamiento desconocidas en los parentales dificultan la observación directa de los genotipos a través de los fenotipos, disminuyendo la fiabilidad del método directo.

En la práctica, la estima de la frecuencia de recombinación a partir de datos fenotípicos implica una serie de técnicas estadísticas, que se basan mayoritariamente en aproximaciones de máxima verosimilitud (Ott, 1991; Liu, 1998; Nuez, 2000).

El método de estimación de máxima verosimilitud fue descrito por primera vez por Fisher (1922a; 1922b). El propósito de este método es buscar a partir de datos observados el valor de un parámetro p , que maximice la probabilidad de observación de dichos datos.

Los supuestos del método de máxima verosimilitud se ajustan a los de la distribución binomial: existen dos variables independientes cuyas probabilidades no cambian con el tiempo. Así, la función de distribución binomial puede servir como expresión de la función de verosimilitud:

$$[n!/(h!(n-h)!)]p^h(1-p)^{n-h}$$

donde n es el tamaño de la muestra, h es el número de sucesos y p es la probabilidad de h .

En el caso de la función de verosimilitud, se ignora la parte de la expresión que no depende del parámetro p , a fin de evitar el problema computacional que podría ocasionar el cálculo de los valores factoriales. Así, la función de verosimilitud se expresaría como:

$$p^h(1-p)^{n-h}$$

En el caso de la estimación de frecuencias de recombinación entre marcadores, las dos variables independientes serían el número de recombinantes y no recombinantes observados en la población y el parámetro a estimar sería la frecuencia de recombinación (r). Puesto que cada gameto tiene una probabilidad r de ser recombinante y una probabilidad $1-r$ de no serlo, la función de verosimilitud se definiría como:

$$L(r) = r^R(1-r)^{N-R}$$

Donde N es el número de productos meióticos informativos, R es el número de meiosis recombinantes y $N-R$ es el número de meiosis no recombinantes.

El método procede ensayando una serie de valores para r , por ejemplo $r = 0,0, 0,1, 0,2, 0,3, 0,4, 0,5$ y computa $L(r)$ para cada uno de ellos, la verosimilitud de los datos será mayor para un determinado valor de r , el cual se toma como la mejor estima de r (\hat{r}). De hecho, la mejor estima de r para cualquier muestra es claramente la proporción de recombinantes observada en la muestra (Figura II.7).

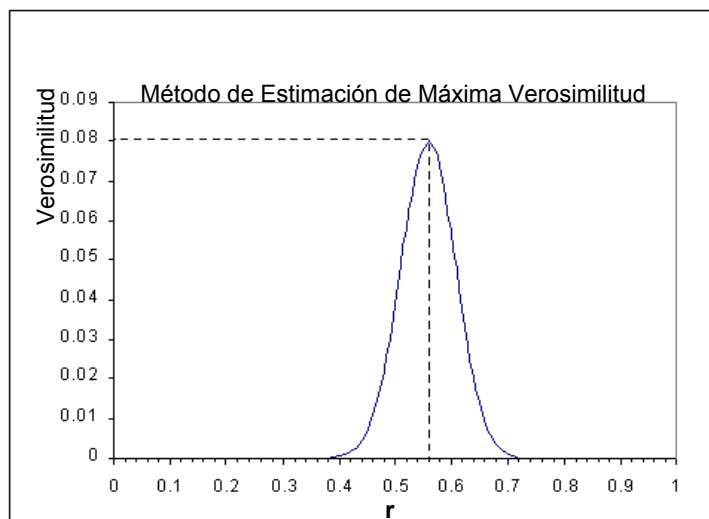


Figura II.7. La mejor estima de r será la proporción de recombinantes observados. Así, para dos loci no ligados la estima de r que maximiza la verosimilitud rondará 0,5.

II.3.3.2. Determinación de ligamiento entre pares de marcadores: Test de razón de verosimilitud y método de puntuación LOD.

En el programa MAPMAKER la determinación de ligamiento entre dos loci se basa en el test de la razón de verosimilitudes (LR). Este test permite comparar la verosimilitud de los datos bajo una hipótesis (hipótesis alternativa) frente a la verosimilitud de los mismos datos bajo una hipótesis más restrictiva (hipótesis nula).

$$LR = L(H_A)/L(H_0)$$

Resuelta la razón de verosimilitudes, los puntos de ventaja de una hipótesis frente a la otra se conocen como *odds*.

Por razones relacionadas con el análisis computacional de los datos, las verosimilitudes se expresan en escala logarítmica, de tal forma que el resultado del logaritmo de la fracción de verosimilitudes se conoce como LOD (logaritmo de *odds*; Barnard, 1949).

$$\log_{10}(LR) = \log_{10}[L(H_A)/L(H_0)] = \log_{10}L(H_A) - \log_{10}L(H_0)$$

En el caso particular del análisis de ligamiento, las dos hipótesis a comparar son ligamiento (H_A , $r < 0,5$) frente a segregación independiente (H_0 , $r = 0,5$). La diferencia entre las verosimilitudes para cada una de las hipótesis se conoce como puntuación LOD ($Z(r)$) o *LOD score*.

$$Z(r) = \log_{10}[L(\hat{r})/L(0,5)]$$

La puntuación LOD es máxima cuando el valor de \hat{r} corresponde al estimado según el método de máxima verosimilitud.

Cuanto menor sea el valor de \hat{r} respecto a 0,5, mayor es la evidencia de ligamiento. Por otro lado, cuando no hay ligamiento ($\hat{r} = 0,5$), la puntuación LOD es igual a 0 ($\log_{10}(1) = 0$). Esta determinación de asociación entre loci a partir de la puntuación LOD es lo que se conoce como método *LOD score* y fue aplicado por primera vez en estudios de mapeo genético en humanos por Morton (1955).

II.3.3.3. Criterios de asignación de ligamiento

Valores de puntuación LOD positivos indican evidencia a favor del ligamiento entre los marcadores, por el contrario, valores iguales a cero o negativos indican evidencias a favor de la independencia de los marcadores. El máximo *LOD score* ($Z_{\text{máx}}$) sirve como una medida del peso de los datos en favor o en contra de la hipótesis de ligamiento. Cuando $Z_{\text{máx}}$ alcanza un cierto valor crítico, se dice que los datos expresan evidencia significativa de ligamiento. Generalmente, este valor crítico es el originalmente propuesto por Morton (1955), $Z = 3,0$. La interpretación de esta puntuación LOD es que la hipótesis alternativa es 10^3 veces más probable que la hipótesis nula (Liu, 1998).

El test de razón de verosimilitudes, predice que asintóticamente, bajo la hipótesis nula de ausencia de ligamiento, $4,6 \times Z_{\text{máx}}$ sigue una distribución χ^2 con un grado de libertad. Consecuentemente, dado el nivel crítico de $4,6 \times 3 = 13,8$ para el χ^2 , el nivel de significación asintótico (α) de este test LR es igual a 0,0002.

La precisión de la estima r se puede valorar a través de un intervalo de confianza específico. Este intervalo se puede definir a partir del trazado de la curva del LOD score para diferentes valores de r (Figura II.8).

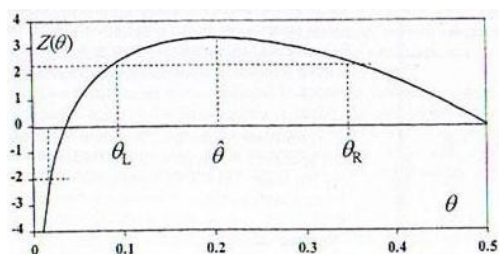


Figura II.8. Identificación del intervalo de confianza de la estima de \hat{r} a partir de la curva del lod score. Según Ott, 1991.

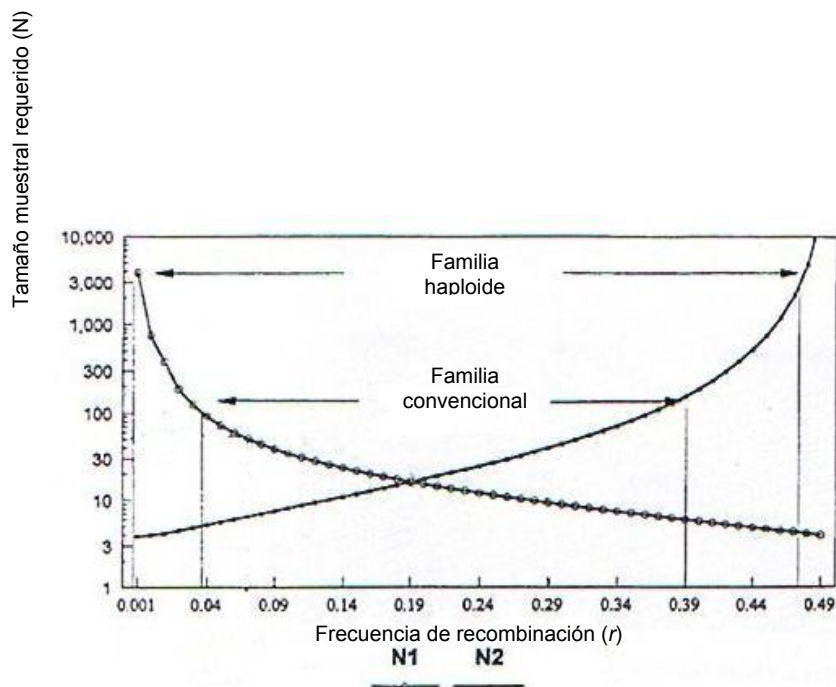
La curva representada llega a su máximo para $Z_{\text{máx}} = 3,348$ y $\hat{r} = 0,20$. Según las recomendaciones de Conneally *et al.* (1985), el intervalo de confianza $Z_{\text{máx}} - 1$ se obtiene dibujando una línea horizontal a una unidad de $Z_{\text{máx}}$ por debajo del LOD score ($Z = 2,348$), siempre y cuando $Z_{\text{máx}} \geq 3$ (de otra forma, no se construiría un intervalo de confianza). Los puntos de intersección de esta línea con la curva del LOD score marcan los extremos derecho e izquierdo del intervalo cuando se proyectan sobre el eje \hat{r} .

Particularmente, el programa MAPMAKER asume por defecto que dos loci están ligados cuando la fracción de recombinación entre ellos es menor de 0,5 (distancia genética máxima 50 cM) y la puntuación LOD mínima es mayor o igual a 3,0.

II.3.3.4. Tamaño muestral y precisión de la estima de r

Una de las principales ventajas de trabajar con material haploide es el elevado número de productos meióticos que pueden ser analizados en un único individuo (Lie *et al.* 1994). Sin embargo, es necesario encontrar un tamaño muestral que consensúe la estimación de niveles significativos de recombinación (r) y las limitaciones de tiempo y coste del análisis.

Lie *et al.* (1994) calcularon los tamaños muestrales mínimos N_1 y N_2 para determinar si una frecuencia estimada de recombinación r difiere significativamente de 0 o de 0,50, respectivamente (Figura II.9). Igualando $N_1 = N_2$ se obtiene el “ r óptimo” que corresponde al punto dónde se cruzan las dos curvas de la Figura II.9. Este punto representa la frecuencia de recombinación ($r = 0,18$) que requiere un menor tamaño muestral ($N = 18$).



En la Figura II.9. Tamaño muestral (N) requerido como función de la frecuencia de recombinación (r). N_1 : tamaño muestral mínimo para distinguir entre una \hat{r} estimada y $r = 0$ (no recombinación). N_2 : tamaño muestral mínimo para distinguir entre una r estimada y $r = 0,50$ (segregación independiente). Estimamos dentro de un intervalo de literatura de mapeo genético en peces, el tamaño muestral más frecuente cuando se trabaja con familias haploides varía entre los 40 individuos (Kocher *et al.*, 1998; Gates *et al.*, 1999; Sun y Liang, 2004) y los 70 (Young *et al.*, 1998; Poompuang y Na-Nakorn, 2004). Considerando la literatura y el gráfico de la Figura II.9, se decidió trabajar con un tamaño muestral de 50 embriones haploides. Según las previsiones de Lie *et al.* (1994), este tamaño muestral permite diferenciar recombinación de ligamiento completo ($r=0$) cuando la

frecuencia de recombinación entre marcadores es $\geq 7\%$ e igualmente, permite distinguir ligamiento de segregación independiente siempre que la frecuencia de recombinación entre los marcadores no supere el 30%.

II.3.4. Ordenación de loci: análisis multipunto

Para testar los $n!/2$ posibles órdenes de n marcadores en un grupo de ligamiento hay dos estrategias posibles: análisis de dos puntos y multipunto, basados en estimas de frecuencias de recombinación bilocus o multilocus, respectivamente. Puesto que el análisis dos puntos sólo considera dos loci cada vez, este análisis no puede indicar si un marcador se localiza a la “derecha” o “izquierda” del otro. Por el contrario, el análisis multipunto es más informativo, ya que permite el ensayo simultáneo de los diferentes marcadores evaluados, considerando la frecuencia de dobles entrecruzamientos. Esto permite establecer entre qué dos marcadores es más probable que se localice un tercero.

II.3.4.1. Análisis de tres puntos

La extensión inmediata del modelo de dos puntos es el ordenamiento de tres puntos y de la misma manera, el ordenamiento multilocus es una extensión del ordenamiento de tres puntos.

Dados tres loci A, B y C, la determinación de su orden en el mapa puede abordarse de distintas maneras, entre ellas, la identificación de los dobles recombinantes y a partir de estos determinar el orden de los loci. Otra aproximación habitual para determinar el orden de los tres loci se basa en la magnitud de las estimas de las frecuencias de recombinación de dos puntos. Si por ejemplo, $\hat{r}_{AB} = 0,1$, $\hat{r}_{AC} = 0,22$, $\hat{r}_{BC} = 0,3$, los loci localizados en los extremos mostrarán el mayor valor de frecuencia de recombinación, siendo B-A-C el orden más probable. El otro método comúnmente usado para el ordenamiento de tres loci es comparar las verosimilitudes de los tres órdenes posibles. El orden con la mayor probabilidad es el orden más verosímil. La aproximación

por máxima verosimilitud es el método de ordenación usado por MAPMAKER (Lander *et al.*, 1987).

II.3.4.2. Estimación de la tasa de recombinación multilocus

La aproximación multipunto asegura que en todas las meiosis estén presentes marcadores flanqueantes informativos.

Durante las últimas décadas, se han desarrollado muchas aproximaciones para construir un modelo multilocus. El método de mínimos cuadrados (Jensen y Jorgensen, 1975a; 1975b), usado en el programa JOINMAP (Stam, 1993); el método de unión de máxima verosimilitud (Zhao *et al.*, 1990); y el EM (Expectation-Maximization) o algoritmo de Lander y Green (Lander y Green, 1987), son algunos de los métodos más usados para estimar tasas de recombinación multilocus.

MAPMAKER usa el algoritmo EM (Dempster *et al.*, 1977; Lander y Green, 1987; Weir, 1996) para computar numéricamente las estimas de recombinación multipunto de máxima verosimilitud. El algoritmo EM es un algoritmo de esperanza/ maximización que procede en dos pasos: el paso E (esperanza) en el que, asumiendo ausencia de interferencia, se usa un valor supuesto de r (r') para computar el número esperado de gametos recombinantes para cada intervalo; y el paso M (maximización) que usando estos valores esperados como si fueran verdaderos computa la estima de máxima verosimilitud (r'') para la frecuencia de recombinación. Este proceso se repite hasta que la nueva estima de la fracción de recombinación iguale a la primera ($r = r'$) de forma que converjan en un valor de máxima verosimilitud. El algoritmo EM maximiza la verosimilitud más sobre los loci que sobre los individuos en un pedigrí, lo cual permite el análisis conjunto de un gran número de loci, pero limita el tamaño de los pedigrís analizables.

II.3.4.3. Ordenación de los loci dentro de sus grupos de ligamiento.

Con n marcadores, el procedimiento ideal para identificar el orden de los loci sería computar la función objetiva para cada uno de los $n!/2$ órdenes posibles y entonces hacer un rango de los órdenes resultantes eligiendo como mejor aquel con una mayor función objetiva. Sin embargo, con un número grande de marcadores resultaría inviable. Es necesario utilizar algoritmos que permitan estimar el mejor orden de marcadores sin considerar todos los órdenes posibles. Existen distintos criterios para definir el mejor orden de los loci, tales como la suma mínima de fracciones de recombinación adyacentes (SARF; Falk, 1989), la suma mínima de puntuaciones LOD adyacentes (SALOD, Weeks y Lange, 1987), la suma mínima de la probabilidad de dobles entrecruzamientos (PDR; Knapp *et al.*, 1989) y la máxima verosimilitud (ML: *Maximum Likelihood*; Lander y Green, 1987).

Los diferentes programas de análisis de ligamiento usan diferentes métodos computacionales o algoritmos basados en alguno de estos criterios de ordenación para resolver la disposición de los loci dentro de un grupo de ligamiento. MAPMAKER ordena los loci aplicando el criterio de máxima verosimilitud (Lander *et al.*, 1987). Dado un grupo de ligamiento, MAPMAKER comienza ordenando un pequeño subgrupo de marcadores informativos, para los que calcula la verosimilitud (L) de todos los órdenes posibles. Entre estos órdenes MAPMAKER selecciona como mejor aquel que maximice el valor de L .

$$\log_{10}L = \log_{10}L(\text{orden } x)/L(\text{mejor orden})$$

El valor del logaritmo de verosimilitudes para el mejor orden es 0,0

El subgrupo de marcadores ordenados constituye un mapa de referencia o *framework* dentro de un grupo de ligamiento de tamaño medio. Este *framework* no contiene más de 5 marcadores informativos. Las posiciones de cada uno de los restantes loci del grupo de ligamiento relativas a las del mapa *framework* se testan sucesivamente. Cada marcador se intenta posicionar en este mapa y la localización más adecuada se toma como su posición estimada en el mapa. De esta forma, un marcador tras otro se van

ordenando en el grupo respecto a su decreciente informatividad. Para los marcadores menos informativos no será posible encontrar una posición precisa en el grupo.

Puesto que en un cruce experimental se dispone de datos informativos de todos los individuos para todos los loci en estudio (excepto para algunos marcadores en los que problemas técnicos impiden la identificación de los genotipos) previamente al análisis multipunto, se puede inferir el orden de los loci usando un análisis de tres puntos, a fin de limitar el número de órdenes posibles. Usando el comando `THREE POINT`, `MAPMAKER` precomputa las probabilidades tres puntos para cualquier grupo de ligamiento. Cuando esta función está activa, otros comandos de ordenación usarán el análisis tres puntos para restringir el conjunto de órdenes de loci que deben ser examinados. Los órdenes que resulten incompatibles con los resultados del análisis tres puntos se excluirán para posteriores análisis, mientras que los órdenes compatibles serán objeto de análisis multipunto completos. Sin embargo, el comando `THREE POINT` puede producir resultados incorrectos, particularmente con marcadores estrechamente ligados y cuando hay errores de genotipado. Así, se recomienda el uso del análisis tres puntos con límites estrictos del logaritmo de la verosimilitud.

Una vez determinado el orden de los marcadores en un grupo de ligamiento, es posible visualizar las distancias de mapa (cM) entre los marcadores ordenados, así como la longitud total del grupo de ligamiento, usando el comando `MAP`. Para calcular estas distancias, `MAPMAKER` usa por defecto la función Haldane. Alternativamente el usuario puede seleccionar la función Kosambi como opción de trabajo. Particularmente, en la construcción del presente mapa genético de rodaballo se utilizó la función Kosambi, considerando la descripción en peces de niveles de interferencia relativamente elevados (Thorgaard *et al.*, 1983; Johnson *et al.*, 1996; Kocher *et al.*, 1998).

Para la representación gráfica de los grupos de ligamiento se utilizó el programa de libre acceso MAPCHART 2.0 (Voorrips, 2002).

II.3.5. Detección de errores de genotipado: algoritmo de Lincoln y Lander (1992).

Los errores en un mapa genético denso tienden a producir sucesos estadísticamente improbables, tales como dobles entrecruzamientos entre loci adyacentes que llevan a incrementar la longitud genética de los intervalos y a interferir en la correcta interpretación del orden de los marcadores (Lincoln y Lander, 1992; Collins *et al.*, 1996; Hackett y Broadfoot, 2003). En los mapas genéticos más densos, se ha estimado que la tasa de error podría estar en el rango de 0,5-3,0% (Dracopoli *et al.*, 1991; Patterson, 1991; Dietrich *et al.*, 1992). Tales tasas de error tienen sólo ligeros efectos en un mapa genético espaciado, sin embargo, sus efectos son considerables cuando el mapa aumenta en densidad.

Los primeros autores en abordar la problemática asociada a la introducción de errores fueron Lincoln y Lander (1992). Posteriormente, otros investigadores desarrollaron distintos métodos estadísticos para la identificación automatizada de los genotipos erróneos (Brzustowicz *et al.*, 1993; Ott, 1993; Ehm *et al.*, 1996; Stringham y Boehnke, 1996; O'Connell y Weeks, 1998; 1999; Douglas *et al.*, 2000). Desafortunadamente, ninguno de los métodos propuestos resulta óptimo para detectar por sí solo todos los errores de genotipado.

El método ideado por Lincoln y Lander e incluido en el programa MAPMAKER, es un método de tolerancia al error, que implica la construcción de mapas genéticos bajo un modelo que permite errores de genotipado. Como resultado, el método puede reconocer cuando un evento es más probable que

sea el resultado de un error que de una recombinación. Esto evita la expansión del mapa e identifica los genotipos a reevaluar.

Conceptualmente, el método parte de la consideración de los datos de genotipado como fenotipos en lugar de genotipos. Así, los genotipos reales serán muy próximos a los fenotipos observados, pero no idénticos debido a los errores. En otras palabras, los genotipos exhibirían penetrancia incompleta. Esta característica puede incluirse en la aproximación tradicional de máxima verosimilitud para la cartografía génica, basta especificar una función de penetrancia. Esta función, basada a su vez en una supuesta tasa de error (ϵ), definiría la probabilidad de observar un cierto fenotipo dado el genotipo real. La función de penetrancia se aplicaría entonces a todos los cálculos multipunto para la construcción del mapa más probable, en el que los posibles errores se identificarían como aquellos datos dónde los genotipos inferidos probabilísticamente no se corresponden con los datos observados.

En el caso concreto de un retrocruzamiento la función de penetrancia se expresaría como:

$$P(\text{genotipo observado} = X \mid \text{genotipo real} = G) = \begin{cases} 1 - \epsilon & \text{si } X = G \\ \epsilon & \text{si } X \neq G \end{cases}$$

En el mapa resultante del análisis multipunto se puede calcular el LOD score de error ($\text{LOD}_{\text{error}}$) asociado a cada genotipo. Siendo X_{ij} el genotipo observado y G_{ij} el genotipo real del individuo i en el locus j , el $\text{LOD}_{\text{error}}$ se definiría como el logaritmo de la razón de probabilidades de que se origine el conjunto de datos X siendo X_{ij} incorrecto, frente a que se origine siendo X_{ij} correcto:

$$\text{LOD}_{\text{error}}(i, j) = \log_{10} [P(X \mid X_{ij} \neq G_{ij}, \hat{r} \wedge \epsilon) / P(X \mid X_{ij} = G_{ij}, \hat{r} \wedge \epsilon)]$$

\hat{r} sería la frecuencia de recombinación estimada en el mapa en curso y ϵ sería la tasa de error supuesta.

Los resultados X_{ij} con los valores más altos de LOD_{error} son los errores potenciales más probables dados los datos. Así, cada mapa puede ser reevaluado tras la corrección de los genotipos en los cuales el LOD_{error} supere un cierto valor crítico fijado por el usuario.

El LOD_{error} puede variar en un rango de -10,0 (o menor) a +10,0 (o mayor), valores superiores indican mayor posibilidad de error en una escala logarítmica. Límites de LOD_{error} entre 0,0 y +2,0 detectan la mayoría de los errores de tipado en los conjuntos de datos típicos.

En este trabajo, los cálculos multipunto se efectuaron suponiendo una tasa de error *a priori* del 1% y se reevaluaron todos los genotipos cuyo LOD_{error} superó 1,0.

II.4. ESTUDIO DE LA DISTRIBUCIÓN DE MARCADORES A LO LARGO DEL GENOMA.

En el mapa femenino de *S. maximus* se testó la aleatoriedad de la distribución de los loci microsatélite y AFLP, tanto dentro de cada grupo de ligamiento como a lo largo del mapa total.

II.4.1. Distribución de marcadores dentro de los grupos de ligamiento.

Cada grupo de ligamiento con más de dos posiciones únicas o *bins* se dividió en intervalos de 10 cM y se contó la frecuencia de intervalos conteniendo desde 0 a 5 marcadores. La frecuencia de cada tipo de intervalo se analizó en cada grupo de ligamiento (LG: "linkage group") considerando dos casos: (i) incluyendo sólo los marcadores *framework* e (ii) incluyendo los *framework* y los marcadores accesorios. En cada caso, las frecuencias observadas se compararon con las esperadas según la distribución de Poisson.

II.4.2. Distribución de marcadores a lo largo del mapa.

La distribución de los marcadores a lo largo del genoma de *S. maximus* se evaluó desde dos perspectivas: (1) considerando todos los marcadores mapeados, independientemente de su composición y (2) teniendo en cuenta el origen y la composición de bases nucleotídicas de los marcadores.

II.4.2.1. Evaluación de la distribución de marcadores independientemente de su composición.

Para evaluar la aleatoriedad de la distribución de marcadores se aplicaron dos pruebas:

- Test de Pearson: a fin de evaluar la correlación entre el tamaño de los grupos de ligamiento y el número de marcadores de cada grupo.
- Comparación de las frecuencias observada y esperada de intervalos con un número n de marcadores: los grupos de ligamiento con al menos tres loci únicos o *bins* se dividieron en intervalos de 10 cM. Se analizó entonces el número de intervalos conteniendo entre 0 y 5 marcadores en el mapa *framework* y en el mapa total (incluyendo los loci *framework* y los marcadores accesorios). La frecuencia observada de cada tipo de intervalo se comparó con la frecuencia esperada según la distribución de Poisson utilizando un test χ^2 a un nivel de significación de 0,05.

II.4.2.2. Evaluación de la distribución de marcadores en función del tipo de marcador y de su composición.

Una vez analizada la distribución global de los marcadores a lo largo del mapa, se evaluó la disposición en el mapa de los marcadores en función de su similitud, ya sea en cuanto a su composición nucleotídica o en cuanto a su origen.

- Marcadores microsatélite. En el caso de los marcadores microsatélite genotipados en la familia H8, no puede establecerse una diferencia entre marcadores en cuanto al tamaño del motivo de repetición, ya que a excepción del Sma1-125INRA, todos los marcadores presentan motivos dinucleotídicos. Sin embargo, puede evaluarse la distribución de estos loci en función de la complejidad de dicho motivo. Se clasificaron entonces en marcadores simples (el motivo de repetición estaría formado por una secuencia dinucleotídica) o complejos (incluirían en la secuencia de repetición algún nucleótido ajeno al motivo). La distribución de ambos tipos de marcadores se estudió a lo largo de los grupos de ligamiento con más de 3 posiciones únicas aplicando un test χ^2 de contingencia siendo $P > 0,05$.

Los mismos marcadores microsatélites fueron también clasificados en función de su origen (1, Coughlan *et al.*, 1996; 2, Estoup *et al.*, 1998; 3, Iyengar *et al.*, 2000; y 4, Pardo *et al.*, 2005) y se analizó la aleatoriedad de la distribución de marcadores aislados a partir de distintas genotecas.

- Marcadores AFLP. En cuanto a la distribución de marcadores AFLP, algunos autores han descrito agrupaciones en los grupos de ligamiento de marcadores AFLP amplificados a partir de una misma combinación de cebadores (Kocher *et al.*, 1998; Pérez *et al.*, 2004). En el mapa de *S. maximus* la distribución de este tipo de marcadores a lo largo de los grupos de ligamiento constituidos por al menos tres posiciones únicas o bins se analizó aplicando un test χ^2 de contingencia ($P > 0,05$).

II. 5. ESTIMA DE LA LONGITUD GENÓMICA Y COBERTURA DEL MAPA DE *Scophthalmus maximus*.

Cuando la adición de nuevos marcadores a un mapa de ligamiento no modifica el tamaño del mapa se puede considerar que se ha cubierto la longitud total del genoma en estudio. En ese caso, la longitud genómica total se obtiene de la suma de todos los intervalos mapeados. Sin embargo, antes de llegar a un mapa completo, se pueden obtener estimas de la longitud genómica a partir de datos de ligamiento parciales. Dichas estimas son útiles a la hora de predecir el número de marcadores necesarios para cubrir un genoma y permiten una primera evaluación de la relación entre distancia física y distancia genética, medida como el número de kilobases (kb) de ADN por cM (Gerber y Rodolphe, 1994).

Existen distintas ecuaciones para estimar la longitud genómica (G_e), entre ellas, una de las más citadas en los proyectos de mapeo es la de Hulbert *et al.* (1988) y una modificación posterior de la misma fórmula propuesta por Chakravarti *et al.* (1991).

La estima de Hulbert *et al.* (1988) se basa en el cálculo de la frecuencia esperada (K) de pares de marcadores ligados que elegidos al azar en un genoma de tamaño G tengan un $LOD \geq Z$.

$$K = \frac{N}{2} \left(\frac{2X}{G} \right)$$

siendo N el número de marcadores posicionados a lo largo del genoma y X la distancia (cM) entre dos loci para los cuales se espera una puntuación $LOD \geq Z$. De esta fórmula se puede deducir la longitud genómica:

$$G_e = N(N-1)X/K$$

Chakravarti *et al.* (1991) sugiere una modificación simple del método de Hulbert. Para un valor Z de puntuación LOD , elige entre los pares de loci con $LOD \geq Z$ el par con el valor más alto de r . A partir de este valor se calcula X como la distancia en cM entre dicho par de marcadores.

Usando MAPMAKER, se puede obtener el valor de X mediante el comando `LOD TABLE`, que permite visualizar todos los pares de marcadores dentro de un

grupo de ligamiento, indicando para cada par la puntuación LOD y la distancia en cM.

Otros estimadores de Ge incluyen en su cálculo la longitud que supondría la distancia desde el último marcador de cada cromosoma a su telómero. Así la fórmula de Fishman *et al.* (2001) calcula Ge como:

$$Ge_F = \sum [(2 \times \text{distancia promedio}) + \text{longitud observada del LG}]$$

En el presente mapa de rodaballo, la longitud genómica se calculó como el promedio entre las estimas de Chakravarti y Fishman. Se utilizaron en estos cálculos únicamente los marcadores *framework*, para evitar una sobreestima del tamaño genómico.

A partir de la longitud genómica observada (Go) y la esperada (Ge), se puede calcular la cobertura genómica (Co) como:

$$Co = Go/Ge$$

Además, se puede hacer una estima de la cobertura genómica esperada (Ce) considerando el número de marcadores mapeados. Para este cálculo se utilizó la fórmula de Lange y Boehnke (1982) según se describe en Cervera *et al.* (2001),

$$Ce = 1 - e^{-XN/1.25Ge}$$

Donde X es la distancia máxima entre dos marcadores *framework* adyacentes; N es el número de marcadores *framework*; y Ge es la longitud genómica estimada considerando los extremos cromosómicos.

La comparación entre la cobertura genómica observada y la esperada es un indicador de la aleatoriedad en la distribución genómica de los marcadores utilizados.

III. RESULTADOS

III.1. MATERIAL GINOGENÉTICO HAPLOIDE

III.1.1. Verificación de ginogénesis

La Figuras III.1 muestra el gel de agarosa para la verificación de ginogénesis en algunos de los descendientes de la familia H8. En la figura se aprecia la amplificación de uno u otro de los alelos maternos, pero en ningún caso se detectan alelos paternos, confirmándose la condición ginogenética de los embriones.

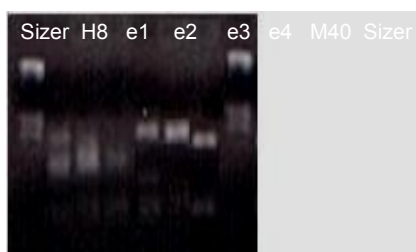


Figura III.1. Verificación de ginogénesis en gel de agarosa al 2%. Se incluyen los productos PCR de la amplificación del microsatélite Sma1-125INRA en la hembra H8 y el macho M40 (carriles 2 y 7, respectivamente) y en 4 de sus descendientes; e1, e2, e3 y e4. Los carriles 1 y 8 del gel fueron cargados con un marcador de tamaño (λ + *Hae*III).

Cuando la diferencia de tamaños entre los alelos implicados es pequeña (entre 2 y 4 pares de bases) resulta difícil identificar en agarosa qué alelos porta la descendencia. En estos casos, la electroforesis vertical en gel de acrilamida seguida de un análisis automatizado de fragmentos resulta imprescindible. Para una mayor precisión en el análisis de fragmentos, todos los productos de amplificación de todos los cruzamientos, independientemente de su tamaño, fueron caracterizados mediante el secuenciador automático *ALF-express* (Amersham Pharmacia Biotech).

El análisis de los loci diagnóstico en las 14 familias sometidas al protocolo para inducción de ginogénesis, demostró el aporte genético exclusivamente materno en los descendientes de las hembras H8, H104, H337, H68, H17, H67, H06, H18 y H47; mientras que en la descendencia de las 5 hembras restantes se encontraron alelos paternos. Estos resultados suponen una tasa de éxito del 64,3% en la inducción de ginogénesis.

III.1.2. Familia de referencia.

A la hora de abordar la construcción de un mapa es importante disponer de familias con un amplio número de descendientes por varias razones. Principalmente, porque un número elevado de descendientes permite obtener estimas más precisas de las frecuencias de recombinación, pero también por cuestiones técnicas. Al trabajar con material embrionario haploide, se debe prever que algunos de estos embriones puedan estar dañados o en condiciones deficientes para obtener una buena amplificación de los marcadores, especialmente de los AFLPs. Por ello, es necesario disponer de un tamaño familiar suficientemente amplio, que permita seleccionar el material en mejor estado. Por esta razón fueron descartadas las familias ginogenéticas con menos de 100 descendientes, las correspondientes a las hembras H337, H67 y H18.

Las progenitoras de las familias ginogenéticas con mayor número de descendientes fueron genotipadas para los 31 marcadores microsatélite publicados hasta entonces en rodaballo, más los cinco microsatélites identificados en el desarrollo de este trabajo (Sma-USC1, Sma-USC2, Sma-USC3, Sma-USC4 y Sma-USC5). La Tabla III.1 recoge los datos de heterozigosis derivados de este análisis.

Tabla III.1. Estimación de la diversidad genética promedio para 36 loci microsatélites (He: heterocigosis esperada; Nei 1987) del parental femenino de las familias ginogenéticas de rodaballo.

Progenitora	Heterozigosis individual (%)
H8	78,6
H379	75,8
H68	72,9
H17	56,6
H6	65,7
H47	74,2

En las hembras analizadas la heterozigosis esperada (He) osciló entre el 56,6% de la H17 y el 78,6% de la H8. Teniendo en cuenta esta información y el tamaño de la progenie disponible, la hembra progenitora más polimórfica

(H8) y 50 de sus descendientes fueron seleccionados como familia de referencia para la construcción del mapa genético de *Scophthalmus maximus*.

III.1.3. Obtención de ADN a partir de embriones haploides. Rendimiento del protocolo de extracción estándar vs. protocolo de extracción optimizado

A fin de cuantificar el rendimiento del protocolo de extracción optimizado, se seleccionaron ocho embriones entre los descendientes de la hembra H8. El ADN de cuatro de ellos fue extraído siguiendo el protocolo estándar de fenol: cloroformo: alcohol isoamílico, mientras que el ADN de los cuatro restantes se extrajo aplicando el protocolo optimizado (apéndices I y II, respectivamente). Los resultados de ambos métodos de extracción se ilustran en la Figura III.2.

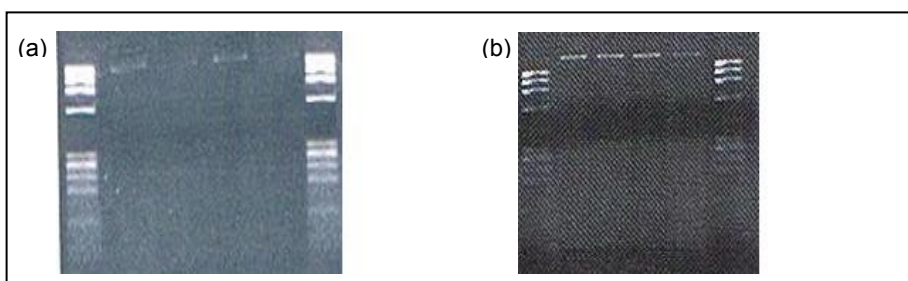


Figura III.2. Electroforesis de ADN genómico de alto peso molecular en geles de agarosa al 1%; (a) ADN genómico de 4 embriones ginogenéticos haploides extraído mediante el protocolo estándar fenol: cloroformo: alcohol isoamílico; (b) ADN genómico de 4 embriones ginogenéticos extraído siguiendo el protocolo optimizado. Los carriles más externos de cada gel corresponden a un marcador de tamaño conocido.

A partir de la migración electroforética, se estimó la cantidad de ADN de alto peso molecular comparando la intensidad de las bandas con las de un marcador de tamaño molecular conocido. Aproximadamente, la concentración de ADN en las muestras extraídas siguiendo el protocolo estándar rondó los 12 ng/μl, mientras que la concentración de ADN obtenido a partir de las muestras en las que se siguió el protocolo optimizado estuvo en torno a los 20 ng/μl.

Puesto que el ADN había sido resuspendido en 20 µl de TE, utilizando el protocolo clásico se obtuvieron unos 240 ng de ADN genómico por muestra, mientras que usando el protocolo optimizado la cantidad de ADN genómico extraído fue de aproximadamente 400 ng. Esto significa un incremento del 40% en el rendimiento de extracción. Estos resultados concuerdan con la descripción comercial del reactivo Phase Lock Gel utilizado en el protocolo optimizado, que asegura un incremento del rendimiento en el proceso de extracción por encima del 30% (Eppendorf, Wetsbury, NY).

III.1.4. Amplificación genómica

Cuando se trabaja con material haploide, la cantidad de ADN de las células es la mitad de la que se podría encontrar en una célula diploide. Por otro lado, la propia condición de haploidía impide un desarrollo normal de las divisiones celulares, limitando la muestras disponibles a embriones de pocos días de desarrollo. Esta limitación en cuanto al ADN de partida no supone un problema para el desarrollo de loci microsatélite, sin embargo, es un factor importante a tener en cuenta cuando interesa caracterizar en un mismo individuo un número elevado de marcadores AFLP y microsatélite.

Bajo esta premisa y a fin de asegurar un stock de ADN por individuo que permitiera abordar el desarrollo de nuevos marcadores AFLP y microsatélite, el ADN extraído con el protocolo optimizado fue sometido a un protocolo de amplificación genómica (GenomiPhi™; Amersham Biosciences). Para cada individuo de la familia H8, a partir de 10 ng de ADN se obtuvieron en torno a 1000-1500 ng de ADN de alto peso molecular según la lectura del espectrofotómetro. La cantidad y calidad del ADN genómico amplificado fueron confirmadas mediante migración electroforética en geles de agarosa teñidos con bromuro de etidio. (Figura III.3).

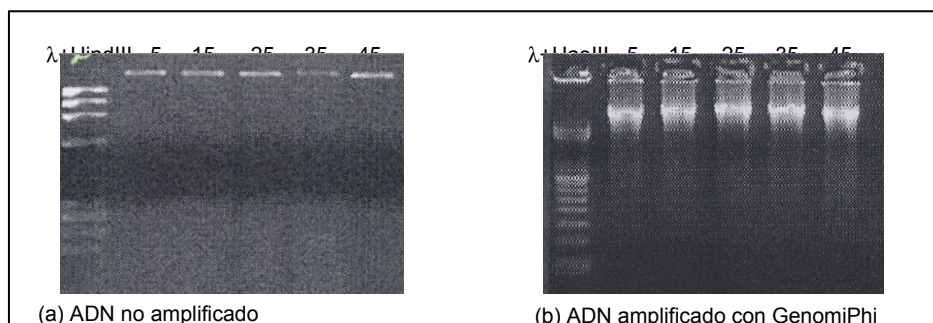


Figura III.3. Verificación en gel de agarosa de la amplificación de ADN genómico procedente de cinco embriones haploides de rodaballo. (a) ADN genómico de alto peso molecular no amplificado. (b) ADN genómico, procedente de los mismos embriones, amplificado con GenomiPhi (Amershan Bioscience). En el primer pocillo de ambos geles se cargaron marcadores de peso molecular conocido (λ + *Hind*III en (a) y λ + *Hae*III en (b)).

III.2. MARCADORES MOLECULARES: Microsatélites y AFLPs

III.2.1. Marcadores microsatélite

III.2.1.1. Nuevos loci microsatélite en *Scophthalmus maximus*.

La búsqueda de nuevos marcadores microsatélite se llevó a cabo en muestras de ADN extraídas de individuos adultos procedentes de poblaciones naturales del rodaballo. En esta búsqueda se siguieron dos estrategias, la amplificación en rodaballo de loci microsatélite caracterizados en otras especies de peces planos (*Pleuronectes platessa* e *Hippoglossus hippoglossus*) y el desarrollo de una librería genómica de rodaballo parcialmente enriquecida en secuencias microsatélite o (STR: *Single Tandem Repeats*).

III.2.1.1.1. Identificación de secuencias microsatélite. Eficiencia de las distintas estrategias de búsqueda.

Amplificación cruzada. La primera estrategia demostró una escasa conservación de los loci microsatélite entre las especies citadas. De 17 pares de cebadores (7 de *P. platessa* y 10 de *H. hippoglossus*), sólo 6 (PL06, PL09, PL115, PL167, Hhi1 y Hhi56) mostraron productos de amplificación que

podieran ser loci microsatélite. Sin embargo, cuando se amplificaron estos loci en una muestra de 5 rodaballos adultos, se observaron amplificaciones inespecíficas y en la mayoría de los casos, de escasa reproducibilidad, en las que no se pudieron identificar los patrones característicos de variación microsatélite. Tan sólo dos de las parejas de cebadores (PL06 y PL167) amplificaron bandas reproducibles entre los individuos de la muestra. Dichos cebadores fueron marcados con fluorescencia (Cy5) y los productos de la amplificación analizados en el secuenciador automático AFL-express. Sin embargo, después del análisis de fragmentos, no se pudo identificar ningún patrón microsatélite entre los productos de la amplificación. Finalmente, se excluyeron todos los pares de cebadores testados y la estrategia de amplificación cruzada fue rechazada por su limitado éxito y por la excesiva inversión que supone, tanto a nivel económico como de trabajo.

Librería enriquecida. A partir de una librería enriquecida en secuencias dinucleotídicas (AC) se purificaron y secuenciaron 78 insertos procedentes de colonias recombinantes. Entre las secuencias amplificadas, 14 se revelaron como potenciales loci microsatélite. La optimización de las condiciones de amplificación de dichas secuencias en dos individuos adultos no relacionados de rodaballo resultó en la identificación de ocho loci microsatélite, que fueron denominados, respectivamente, Sma-USC1, Sma-USC2, Sma-USC3, Sma-USC4, Sma-USC5, Sma-USC6, Sma-USC7 y Sma-USC8. Los seis restantes fueron descartados por la escasa calidad de sus amplificaciones. La Tabla III.2 recoge las condiciones de amplificación PCR para los nuevos microsatélites de rodaballo (Pardo *et al.*, 2005).

Tabla III.2. Condiciones de amplificación de ocho nuevos microsatélite de *S. maximus* caracterizados en el presente estudio.

Locus	Ta (°C)	MgCl ₂
Sma-USC1	62	1,5
Sma-USC2	62	1,5
Sma-USC3	55	1,5
Sma-USC4	55	1,5
Sma-USC5	50	1,5
Sma-USC6	58	1,5
Sma-USC7	58	1,5
Sma-USC8	62	1,5

III.2.1.1.2. Diversidad genética de los nuevos microsatélites

El análisis de variabilidad alélica de los ocho nuevos microsatélite reveló tres loci monomórficos en una muestra poblacional salvaje (Sma-USC6, Sma-USC7 y Sma-USC8,) y posiblemente, no informativos en la familia de referencia. La Tabla III.3 recoge los datos referentes al análisis de variabilidad de los cinco marcadores polimórficos.

Tabla III.3. Características de los cinco nuevos marcadores microsatélite de *S. maximus*. Para cada locus se indica el rango alélico en pares de bases (pb), número de alelos, heterozigosis observada (Ho) y esperada (He), y significación del test de ajuste al equilibrio Hardy-Weinberg (valor P; $P < 0.05$), así como su código de acceso a GenBank. En la última fila, se muestra la media poblacional para los estimadores de diversidad genética utilizados.

Locus	Rango de tamaños (pb)	Nº de alelos	Ho	He	Valor P	Nº de acceso
Sma-USC1	114-142	8	0,773	0,844	0,369	AY641451
Sma-USC2	108-148	15	0,636	0,909	0,014*	AY641450
Sma-USC3	206-210	2	0,273	0,304	1	AY641453
Sma-USC4	198-208	2	0,136	0,130	0,538	AY641452
Sma-USC5	205-229	5	0,773	0,650	0,534	AY697865
Media		6,4	0,510	0,575		

La aplicación de un test exacto para valorar el ajuste al equilibrio Hardy-Weinberg (HWE) confirmó un déficit de heterozigotos significativo en el locus Sma-USC2, lo que podría ser indicativo de la presencia de alelos nulos. Igualmente se aplicó un test de desequilibrio de ligamiento genotípico entre todos los pares de loci, que detectó un alejamiento significativo de la hipótesis nula entre los loci Sma-USC3 y Sma-USC4 ($P=0,010$). Ninguno de los dos test resultaron significativos después de aplicar la corrección de Bonferroni ($\alpha=0,007$ y $\alpha=0,02$ para HWE y desequilibrio de ligamiento, respectivamente).

III.2.1.2. Variabilidad de los marcadores microsatélite en la familia de referencia.

Los 31 loci microsatélite de rodaballo recogidos en la bibliografía (Coughlan *et al.*, 1996; Estoup *et al.*, 1998; Iyengar *et al.*, 2000; Bouza *et al.*, 2002; Tabla II.2) más los 5 nuevos marcadores microsatélite identificados en este trabajo fueron genotipados en el progenitor de la familia de referencia (hembra H8). Dicha hembra resultó heterocigótica para 28 marcadores microsatélite y homocigótica para otros 8 marcadores (Sma-03, B11-I 12/6/3, B12-I GT14, 4/3COMP, 1/4AC18, Sma-USC3, Sma-USC4 y Sma-USC5). Por otro lado, 4 de los marcadores microsatélite (Sma3-10INRA, Sma5-111INRA, B18-IIICA70 y 3/20CA17) fueron descartados para el análisis genotípico debido a la escasa reproducibilidad de sus amplificaciones a partir del ADN procedente de los embriones haploides. Los mismos marcadores ya habían demostrado gran cantidad de bandas espúreas en las amplificaciones del ADN diploide de los individuos adultos.

Finalmente, 24 loci heterocigóticos en la hembra H8 fueron amplificados en 50 de sus descendientes haploides. Las amplificaciones se llevaron a cabo utilizando como molde ADN genómico resultante de la amplificación con el kit GenomiPhi™ (Amersham Biosciences).

III.2.1.3 Relación de ligamiento genético entre loci microsatélite.

Una vez completado el genotipado de la familia H8 (la hembra más su descendencia haploide) se utilizaron los datos para analizar las posibles relaciones de ligamiento genético entre los marcadores microsatélite. De los 24 marcadores analizados, 9 microsatélites resultaron no ligados, mientras que los 15 restantes mostraron relaciones de ligamiento entre pares de loci (Tabla III.4). Estos 15 marcadores se distribuyeron en 5 dobletes formados cada uno por 2 marcadores ligados y un grupo de ligamiento constituido por cinco marcadores (Figura III.4).

Tabla III.4. Ligamientos significativos entre pares de loci microsatélite en *S. maximus*. Para cada par de marcadores ligados se incluye la frecuencia de recombinación y la puntuación LOD del ligamiento

Pares de marcadores ligados	Frecuencia de recombinación	Puntuación LOD
Sma3-12INRA - 3/3GTr	16,9	5,28
Sma3-129INRA - Smax-02r	12,8	5,50
2/5TG14r - 5/4CA20	0,0	14,75
2/5TG14r - F1-OCA19	11,3	6,73
F1-OCA19 - Smax-04b	17,3	4,42
F1-OCA19 - Smax-04r	17,3	4,42
Smax-04b - Smax-04r	0,0	13,8
2/5TG14r - Smax-04b	18,6	4,40
2/5TG14r - Smax-04r	18,6	4,40
5/4CA20 - Smax-04b	18,6	4,40
5/4CA20 - Smax-04r	18,6	4,40
4/4CA4/13 - 3/9CA15	8,7	6,09
F12-I AG18 - F12-I TGr	0,0	13,55
F8-ICA26r - F8-I 11/8/17	0,0	13,55

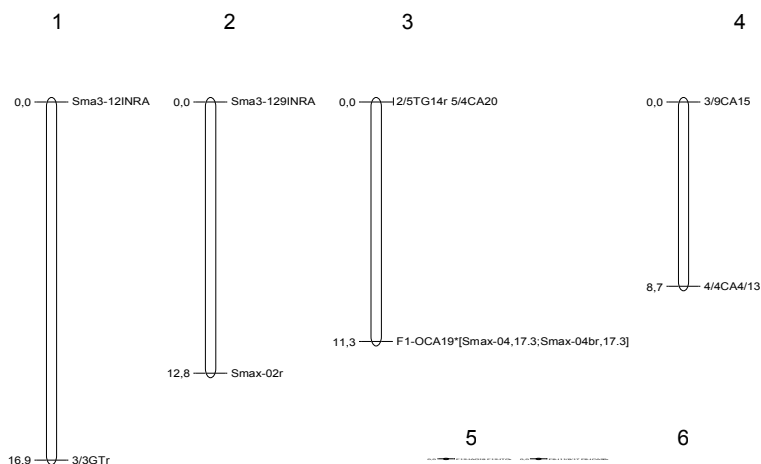


Figura III.4. Grupos de ligamiento genético para 15 marcadores microsatélite de *S. maximus*. Las distancias genéticas se representan en el lado izquierdo de los grupos y vienen dadas en cM asumiendo la función de mapa Kosambi. En el grupo 3, los marcadores Smax-04 y Smax04b no pudieron posicionarse con exactitud y se muestran como accesorios al lado del locus más próximo, indicando a la derecha de cada marcador la distancia de mapa que los separa de dicho locus. Se indica con un asterisco los marcadores que demostraron desviación de la segregación mendeliana.

Los grupos de ligamiento 5 y 6 están formados por pares de marcadores completamente ligados. Los pares de marcadores implicados fueron aislados a partir de un mismo clon y la distancia entre los microsatélite ligados de cada par fue inferior a 2,3 kb (Iyengar *et al.*, 2000). La misma situación de ligamiento completo se observó en dos pares de marcadores situados en el grupo 3: el par Smax-04 y Smax-04b y el par 2/5TG14 y 5/4CA20. Los primeros pertenecían a un mismo clon y se encuentran a unos pocos pb de distancia (Coughlan *et al.* 1996; Bouza *et al.* 2002); los segundos fueron descritos por Iyengar *et al.* (2000) como marcadores independientes, sin embargo corresponden a un única posición en el mapa. Cada uno de estos pares de loci se consideraron como un solo *bin* (posición única) a efectos de mapeo.

III.2.2 Marcadores AFLP

III.2.2.1. ADN genómico

El protocolo AFLP se puso a punto en muestras de ADN de cinco embriones haploides de la familia H8. En cada individuo se obtuvieron perfiles AFLP a partir de 200 ng de ADN amplificado con el kit GenomiPhi™ y a partir de la misma cantidad de ADN original extraído según un protocolo optimizado fenol: cloroformo: alcohol isoamílico. Cuando se compararon los perfiles obtenidos se encontró que en un mismo individuo el patrón AFLP resultante de ADN amplificado con GenomiPhi™ presentaba bandas extra respecto al patrón resultante del ADN original (Figura III.5). Estos fragmentos podrían derivar de una degradación parcial del ADN original, que se revelaría en la amplificación genómica. A fin de evitar errores genotípicos derivados de estas circunstancias, el ADN amplificado con el método GenomiPhi™ se descartó para la caracterización de marcadores AFLP.

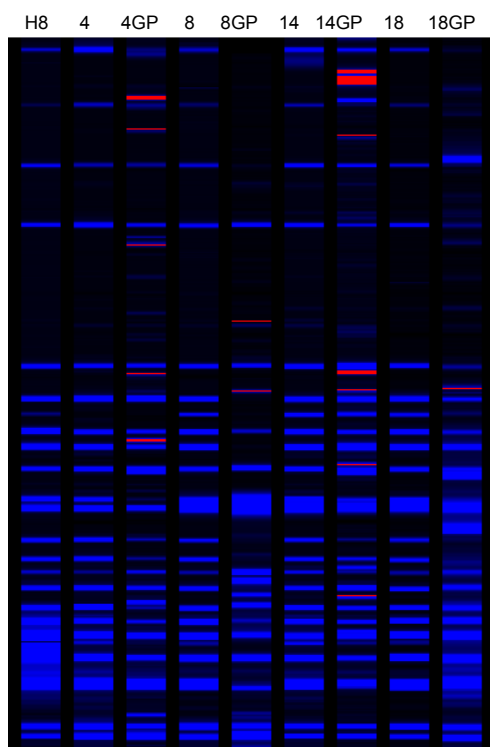


Figura III.5. Perfiles AFLP de la hembra H8 y cuatro de sus embriones haploides (4, 8, 14 y 18) para una misma combinación de cebadores. Para cada embrión se presentan los productos AFLP obtenidos a partir de ADN genómico original y amplificado por GenomiPhi™ (GP). Las bandas rojas representan las bandas no comunes entre los dos perfiles de un mismo individuo. Ninguna de estas bandas aparece en el perfil de la hembra H8.

III.2.2.2 Perfiles AFLP

Los fragmentos resultantes de la digestión enzimática fueron amplificados usando un cebador fluorescente (Figura III.6a). El análisis mediante electroforesis capilar de los productos marcados reveló tamaños moleculares entre los 50 y 650 pares de bases para los productos PCR de la mayor parte de las combinaciones de cebadores. Los niveles más altos de fluorescencia emitidos por estos fragmentos alcanzaron intensidades entre 100.000-150.000 cnts (Figura III.6a). Estos picos de mayor intensidad corresponden a los fragmentos de ADN más cortos, que son los más abundantes tras la digestión enzimática y resultan fácilmente amplificables por la *Taq* polimerasa. La señal de fluorescencia de los fragmentos de tamaño medio varió entre 2000-5000 cnts, pero también se encontraron valores de intensidad más bajos, especialmente entre los fragmentos de ADN de mayor tamaño.

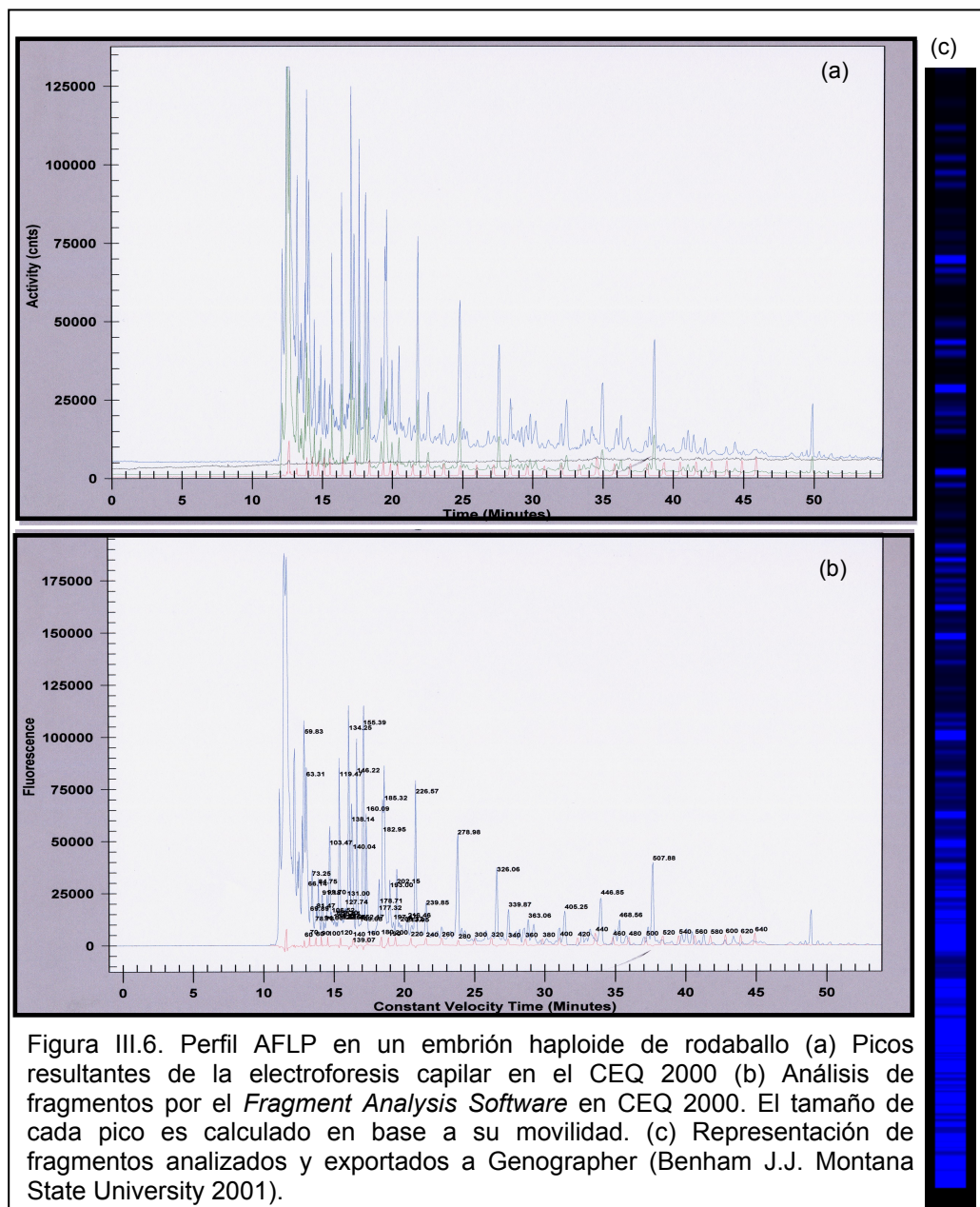


Figura III.6. Perfil AFLP en un embrión haploide de rodaballo (a) Picos resultantes de la electroforesis capilar en el CEQ 2000 (b) Análisis de fragmentos por el *Fragment Analysis Software* en CEQ 2000. El tamaño de cada pico es calculado en base a su movilidad. (c) Representación de fragmentos analizados y exportados a Genographer (Benham J.J. Montana State University 2001).

III.2.2.3. Productividad de los cebadores AFLP

En base al número de picos producidos, la calidad de los mismos y los niveles de polimorfismo, se seleccionaron 17 combinaciones de cebadores entre las 22 testadas en la hembra H8 y 20 de sus descendientes (Tabla III.5).

Con estas 17 combinaciones se genotiparon los 30 embriones restantes para completar la familia H8. Los otros 5 pares de cebadores fueron excluidos bien por la baja calidad de las amplificaciones resultantes o bien por la ausencia o escasez de bandas polimórficas.

Tabla III.5. Combinaciones de cebadores testadas y seleccionadas para el análisis AFLP

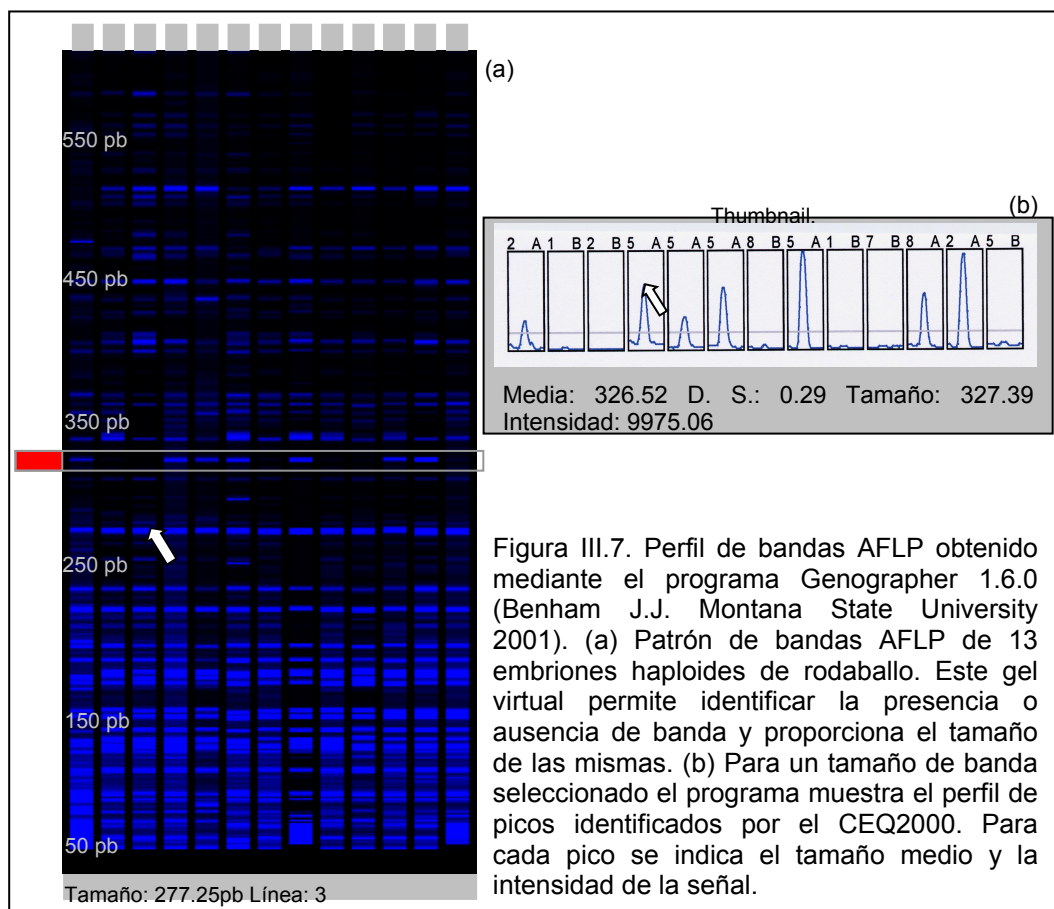
Bases selectivas del cebador <i>EcoRI</i>	Bases selectivas del cebador <i>TaqI</i>							
	T32	T33	T37	T38	T49	T50	T54	T58
	E32	NT	NT	+	NT	+	+	NT
	E33	NT	-	+	+	+	NT	+
	E40	+	-	+	NT	-	+	NT
	E42	+	+	+	NT	+	-	NT

+ combinaciones de cebadores seleccionadas para el análisis AFLP.

- combinaciones de cebadores descartadas.

NT: combinaciones de cebadores no testadas.

Las 17 parejas de cebadores seleccionadas rindieron un total de 906 bandas, variando el número de bandas producidas por combinación de cebadores desde 36 (E42/T33) a 90 (E40/T50). Como promedio, cada combinación de cebadores produjo 53,3 bandas. El tamaño de los fragmentos generados varió desde 50 a 650 pb, si bien, la mayoría de los fragmentos no alcanzó las 450 pb. La Figura III.7a es un ejemplo del patrón AFLP generado por Genographer 1.6.0 (Benham J.J., Montana State University, 2001) para la combinación de cebadores E40/T37 en 13 individuos haploides. En dicho patrón, se detectaron 61 bandas AFLP dentro del rango de 70 a 550 pb.



III.2.2.4. Nivel de polimorfismo detectado

De las 906 bandas totales, 186 (20,5%) fueron identificadas como variables, sin embargo, no todos estas bandas fueron seleccionadas como marcadores para la construcción del mapa. Entre las 186 bandas variables, 54 (29,0%) fueron descartadas por la escasa calidad y la baja reproducibilidad de las amplificaciones. Estas bandas no fueron consideradas en el cómputo de bandas totales, ni en el cálculo de los porcentajes de polimorfismo. Así, el nivel

medio de polimorfismo entre las combinaciones de cebadores analizadas disminuyó a un 15,5% variando entre 5,0% (E32/T54) y 35,1% (E42/T37) (Tabla III.6).

Tabla III.6. Número de bandas totales y número y porcentaje de bandas polimórficas producidas por 17 combinaciones de cebadores AFLP en la familia de referencia.

Combinaciones de cebadores (Keygene)	Extensiones selectivas		Nº bandas totales*	Nº bandas polimórficas	Polimorfismo (%)
	cebador EcoRI	cebador TaqI			
E32/T37	eAAC	tACG	51(7)	8	15,6
E32/T50	eAAC	tCAT	49 (0)	7	14,2
E32/T49	eAAC	tCAG	42(3)	8	19,0
E32/T54	eAAC	tCCT	40(1)	2	5,0
E33/T37	eAAG	tACG	53(5)	11	20,7
E33/T49	eAAG	tCAG	51(1)	8	15,7
E33/T54	eAAG	tCCT	43 (0)	4	9,3
E33/T38	eAAG	tACT	54(2)	9	16,6
E33/T58	eAAG	tCGT	39(9)	12	30,7
E40/T32	eAGC	tAAC	62(7)	8	12,9
E40/T37	eAGC	tACG	55(6)	6	10,9
E40/T50	eAGC	tCAT	88(2)	12	13,6
E40/T54	eAGC	tCCT	50(3)	5	10,0
E42/T32	eAGT	tAAC	42(3)	6	14,3
E42/T33	eAGT	tAAG	34(2)	4	11,7
E42/T37	eAGT	tACG	37(3)	13	35,1
E42/T49	eAGT	tCAG	62(0)	9	14,5
Total/promedio:			852(54)	132	15,5

*entre paréntesis se incluyen el nº de bandas polimórficas que fueron descartadas por su baja calidad de amplificación.

III.2.2.5 Relación entre composición nucleotídica y productividad de los cebadores

Como se observa en la Tabla III.6, las distintas combinaciones de cebadores AFLP rindieron distintos niveles de polimorfismo. Las combinaciones E32/T37, E32/T49, E33/T37, E33/T49, E33/T38, E33/T58 y E42/T37 superaron el 15% de bandas polimórficas, mientras las menos polimórficas apenas llegaron al 10% (E32/T54, E33/T54 y E40/T54).

Estas diferencias en los niveles de polimorfismo no parecen estar relacionadas con la composición de bases de la extensión selectiva en los cebadores, concretamente, con el nivel de GC de estas extensiones

(coeficiente de correlación, $R = -0,04$, $P = 0,85$). Sin embargo, cuando se consideró el número total de bandas amplificado por las 17 combinaciones de cebadores, se observó una tendencia a un incremento en el número de bandas totales a medida que aumentaba el contenido en GC de las bases selectivas del cebador, si bien, tampoco esta correlación resultó significativa ($R = 0,22$, $P = 0,42$).

III.2.3. Caracterización de las meiosis informativas. Análisis de distorsión de la segregación

Entre los marcadores microsatélite, tan sólo uno demostró una desviación significativa de la segregación mendeliana esperada siendo ($0,01 < P < 0,05$; Tabla III.7). Esto supone un 4,1% de los loci analizados, porcentaje que no supera el 5% esperado por azar.

En cuanto a los marcadores AFLP, entre los 132 fragmentos polimórficos seleccionados, 34 (25,7%) mostraron desviaciones de la relación 1:1 esperada bajo la suposición de segregación mendeliana ($P < 0,05$). Entre ellos, 20 fragmentos (15,1%) demostraron una probabilidad $P < 0,01$. Estos fragmentos fueron eliminados para la construcción del mapa, a fin de evitar errores en el análisis de ligamiento. Los 14 marcadores restantes, que demostraron ratios de segregación distorsionadas con valores de significación moderados ($0,01 < P < 0,05$), se incluyeron en el análisis de ligamiento, si bien, no fueron considerados para la construcción del mapa *framework*. En la mayoría de los casos, esta distorsión se debió a un exceso del fenotipo presencia de banda frente a ausencia, a excepción de 7 loci (E33/T54.253, E40/T50.145, E40/T50.241, E42/T33.283, E42/T49.508, E32/T37.284, E33/T37.154, E33/T38.222) donde el fenotipo ausencia de banda prevaleció sobre el de presencia (Tabla III.7).

En total, 136 marcadores (112 AFLPs y 24 microsatélites) fueron incluidos en la construcción del mapa genético de rodaballo. Los marcadores con segregación distorsionada ($0,01 < P < 0,05$), tanto AFLP como microsatélite,

se incluyeron en el mapa una vez establecido el mapa *framework* y su condición de marcadores distorsionados viene indicada por un asterisco (*) en Figura III.11).

Tabla III.7. Distribución de marcadores desviados significativamente de la proporción 1:1 esperada según la segregación mendeliana. Entre paréntesis se indica el grupo de ligamiento al que pertenecen los loci que han sido mapeados.

Locus	Falta de datos ^a	Frecuencia esperada ^b	Frecuencia observada ^b	chi ²
E32/T37.284	11	19.5:19.5	13:26	4,3*
E32/T37.288	11	19.5:19.5	28:11	7,4**
E32/T49.85	7	21,5:21,5	33:10	12,3**
E32/T49.335	4	23:23	39:7	22,2**
E32/T49.421	5	22,5:22,5	40:5	27,2**
E32/T50.274 (LG11)	4	23:23	30:16	4,2*
E32/T50.391	3	23,5:23,5	42:5	29,1**
E33/T37.154	1	24.5:24.5	15:34	7,3**
E33/T37.157	1	24.5:24.5	34:15	7,3**
E33/T37.288	1	24.5:24.5	35:14	9**
E33/T37.466	12	19:19	28:10	8,5**
E33/T38.72	8	21:21	36:6	21,4**
E33/T38.90 (LG5)	10	20:20	27:13	4,9*
E33/T38.94 (LG5)	7	21.5:21.5	29:14	5,2*
E33/T38.222	6	22:22	13:31	7,4**
E33/T49.97	4	23:23	32:14	7,0**
E33/T49.368	4	23:23	39:7	22,2**
E33/T54.253 (LG3)	5	22.5:22.5	15:30	5,0*
E33/T58.137 (LG8)	16	17:17	23:11	4,2*
E40/T32.380	0	25:25	39:11	15,7**
E40/T37.326 (LG8)	11	19.5:19.5	26:13	4,3*
E40/T50.145 (LG4)	5	22.5:22.5	15:30	5,0*
E40/T50.241 (LG11)	3	23.5:23.5	16:31	4,8*
E42/T32.111	4	23:23	31:15	5,5*
E42/T32.294 (LG1)	2	24:24	31:17	4,1*
E42/T33.283	1	24.5:24.5	15:34	7,4**
E42/T33.344	1	24.5:24.5	38:11	14,8**
E42/T37.429	6	22:22	38:6	23,3**
E42/T37.440	7	21.5:21.5	33:10	12,3**
E42/T37.252 (LG1)	2	24:24	31:17	4,1*
E42/T49.150 (LG4)	5	22.5:22.5	30:15	5,0*
E42/T49.152	5	22.5:22.5	38:7	21,3**
E42/T49.505 (LG17)	4	23:23	31:15	5,5*
E42/T49.508	4	23:23	14:32	7,0**
F1-OCA19 (LG3)	4	23:23	31:15	5,6*

*P<0,05; **P<0,01.

^aNumero de individuos que no amplificaron ningún producto para ese cebador.

^bAFLP (presencia : ausencia de banda); microsatélites (alelo A : alelo B).

III.2.4. Identificación de fuentes de error de genotipado en marcadores AFLP

Como ya se comentó en Material y Métodos, la comprobación estadística del ajuste de la segregación alélica a las proporciones mendelianas esperadas, puede considerarse también una herramienta para verificar el genotipado de los individuos. Sin embargo, se ha estimado que un 25% de los errores de genotipado escapan a este test, ya que por azar pueden ajustarse a la segregación mendeliana (error tipo II; Douglas *et al.*, 2002). Concretamente, se han descrito dos casos de incidencias técnicas que afectan particularmente a la distorsión de la segregación mendeliana y que pueden llevar a errores de genotipado: el emparejamiento inespecífico de los cebadores (*mismatch*), con la posible amplificación de una misma banda por más de una pareja de cebadores; y la amplificación por una única pareja de cebadores de fragmentos no homólogos coincidentes en tamaño, fenómeno conocido como homoplasia. Ambas situaciones se han analizado en el presente estudio y los resultados obtenidos se exponen a continuación.

III. 2.4.1. Identificación de amplificaciones inespecíficas

A fin de comprobar la presencia de amplificaciones originadas por emparejamientos inespecíficos de cebadores, se compararon los tamaños de los 186 polimorfismos AFLP generados por las 17 combinaciones de cebadores. Las comparaciones se realizaron entre los productos de amplificación de una combinación de cebadores dada y los originados por los 16 pares de cebadores restantes. De las 289 parejas comparadas, 49 mostraron coincidencia en el tamaño de alguna de las bandas polimórficas amplificadas (Figura III.8). El porcentaje de polimorfismos coincidentes varió desde el 4,5% hasta el 33,3%, siendo el porcentaje medio de coincidencias de 9,8%.

El cómputo del porcentaje medio de polimorfismos comunes resultante de combinaciones de cebadores parecidas (diferenciadas en tan sólo una

base) y el resultante de combinaciones de cebadores bien diferenciados resultó en un valor medio del 12,5% y 9,6%, respectivamente.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1		-	-	(*)	-	-	-	-	-	(*)	-	-	-	-	-	-	-
2			(*)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3				-	-	-	-	-	-	-	-	-	-	*	*	-	-
4					-	*	-	*	*	-	*	-	-	-	-	-	-
5						-	-	*	*	-	-	*	-	-	-	-	-
6							-	-	-	-	*	-	-	-	-	-	-
7								*	-	-	-	-	-	-	-	-	-
8									*	-	*	-	-	-	-	-	-
9										-	-	-	-	-	-	-	-
10											*	-	*	-	(*)	*	-
11												-	-	-	-	*	*
12													-	-	-	*	-
13														-	*	-	-
14															-	-	-
15																-	-
16																	-
17																	

Figura III.8. Parejas de cebadores comparadas para la identificación de polimorfismos coincidentes. Se indica con un asterisco las combinaciones que amplificaron bandas polimórficas de igual tamaño. El asterisco entre paréntesis identifica a las parejas de cebadores que difieren entre sí en tan solo una base nucleotídica. Los números se refieren a las combinación de cebadores: 1: E32/T37; 2: E32/T50; 3: E32/T49; 4: E33/T37; 5: E33/T49; 6: E33/T54; 7: E33/T38; 8: E33/T58; 9: E40/T32; 10: E40/T37; 11: E40/T50; 12: E40/T54; 13: E42/T32; 14: E42/T33; 15: E42/T37; 16: E42/T49; 17: E32/T54.

Como ya se comentó anteriormente, las amplificaciones inespecíficas originan bandas AFLP poco definidas, con amplios *smears* y bajas intensidades de fluorescencia. Teniendo en cuenta las condiciones de calidad consideradas a la hora de seleccionar los marcadores AFLP, es de esperar que muchos de los polimorfismos resultantes de emparejamientos inespecíficos sean descartados en dicho proceso de selección. En el presente trabajo, 74 polimorfismos AFLP fueron descartados para el análisis de ligamiento atendiendo a criterios de calidad de banda y al ajuste a las proporciones de segregación mendeliana. Cuando se evaluó entre los marcadores coincidentes el porcentaje de polimorfismos eliminados *a priori*

atendiendo a criterios de calidad, se apreció una clara diferencia entre los marcadores coincidentes amplificados por parejas de cebadores bien diferenciadas y aquellos amplificados por parejas de cebadores parecidas. En el primer caso, el 28,5% de los polimorfismos coincidentes fueron descartados para el análisis de ligamiento, en el segundo caso, el porcentaje de polimorfismos descartados ascendió al 66,6%.

III.2.4.2. Identificación de bandas complejas.

La presencia de bandas complejas, constituidas por fragmentos de igual tamaño procedentes de distintas regiones genómicas, se testó entre los productos amplificados por la combinación de cebadores E33/T49. Para ello, el patrón de bandas originado por esta pareja de 6 bases selectivas, se comparó con los perfiles originados por 4 parejas de cebadores idénticas a la combinación E33/T49, pero portadoras de una séptima base selectiva: E33/T49+A, E33/T49+T, E33/T49+C y E33/T49+G.

Particularmente, el perfil E33/T49 mostró 51 bandas, de las cuales 8 fueron polimórficas. Entre las bandas polimórficas, dos presentaron distorsión de la segregación mendeliana 1:1 ($P < 0,01$). Cuando se analizó la composición de las bandas polimórficas comparando su patrón con los resultantes de las cuatro combinaciones E33/T49+N, se encontró que cada banda era amplificada por sólo una de las combinaciones con cuatro bases selectivas, con la excepción de una de las bandas con distorsión de segregación, la E33/T49.368, que era amplificada por las combinaciones E33/T49+A y E33/T49+C (Tabla III.8).

Tabla III.8. Identificación de bandas complejas. Para cada uno de los fragmentos amplificados por la pareja de cebadores E33/T49, se indica qué combinación de cebadores E33/T49+N amplifica la misma banda y la condición de la misma como monomórfica (M) o polimórfica (P).

Bandas identificadas	E33/T49	E33/T49+C	E33/T49+G	E33/T49+T	E33/T49+A
E33/T49.080	M	M			
E33/T49.085	M	M			
E33/T49.091	M			M	
E33/T49.097	P			P	

E33/T49.113	M		P	M	
E33/T49.119	M		M	M	
E33/T49.124	M		M		P
E33/T49.128	P				P
E33/T49.132	M	M			
E33/T49.137	M	P		M	
E33/T49.140	M	M			
E33/T49.144	M				M

Tabla III.8. Continuación.

Bandas identificadas	E33/T49	E33/T49+C	E33/T49+G	E33/T49+T	E33/T49+A
E33/T49.146	M	M		P	M
E33/T49.160	M			M	
E33/T49.172	M				M
E33/T49.175	M		M		
E33/T49.186	M			M	
E33/T49.192	M	P		M	
E33/T49.197	P	P			
E33/T49.199	P	P			
E33/T49.201	P	P			
E33/T49.211	P	P			
E33/T49.216	P	P			
E33/T49.236	M		M		
E33/T49.240	M			M	
E33/T49.247	M	M			
E33/T49.250	M	M			
E33/T49.257	M				M
E33/T49.267	M			M	
E33/T49.278	M		M		P
E33/T49.296	M				M
E33/T49.308	M	M	M		
E33/T49.328	M	M			
E33/T49.341	M	M		P	
E33/T49.349	M	M		M	
E33/T49.354	M		M		
E33/T49.362	M			M	
E33/T49.368	P	P			M
E33/T49.381	M	M		P	
E33/T49.388	M	M		P	
E33/T49.424	M	M			
E33/T49.433	M			M	
E33/T49.449	M		M		
E33/T49.470	M			M	
E33/T49.480	M		M		
E33/T49.488	M	M		M	
E33/T49.521	M				M
E33/T49.542	M		M		
E33/T49.559	M				M
E33/T49.566	M	M			
E33/T49.576	M	M			

La banda E33/T49+C.368 es una banda polimórfica cuya segregación alélica se ajusta a la segregación mendeliana 1:1, mientras que la banda E33/T49+A.368 es monomórfica. Los fenotipos positivos del perfil E33/T49+C.368 se correspondieron con los del E33/T49.368, sin embargo éste

último presentó un exceso de positivos (presencia de banda) que no encontraron correspondencia en los fenotipos negativos (ausencia de banda) del E33/T49+C.368. Estos positivos “extra” podrían resultar de la amplificación secundaria de una banda E33/T49+A.368 monomórfica que se evidenciaría en los individuos E33/T49+C.368 negativos (Figura III.9).

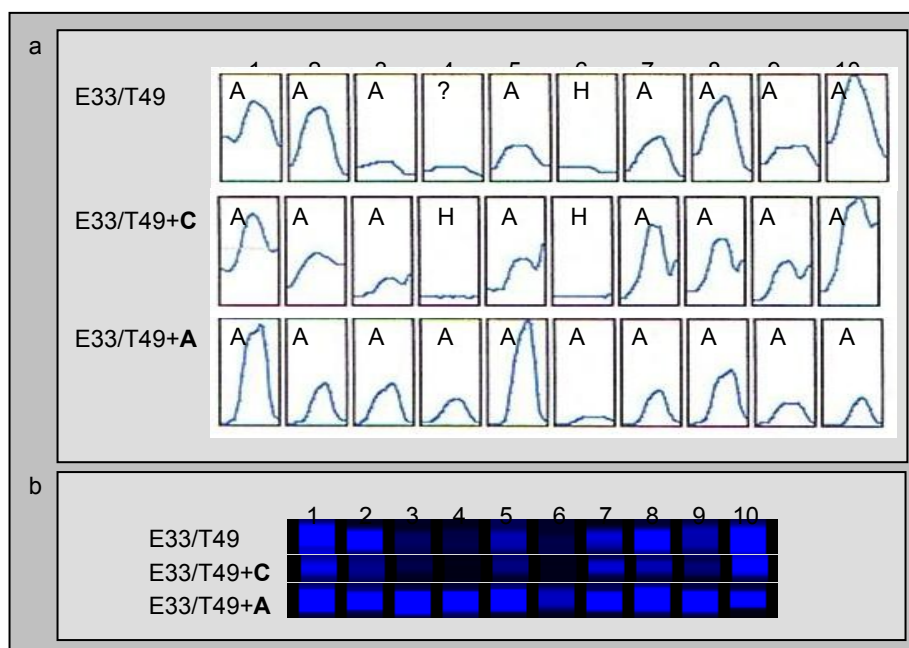
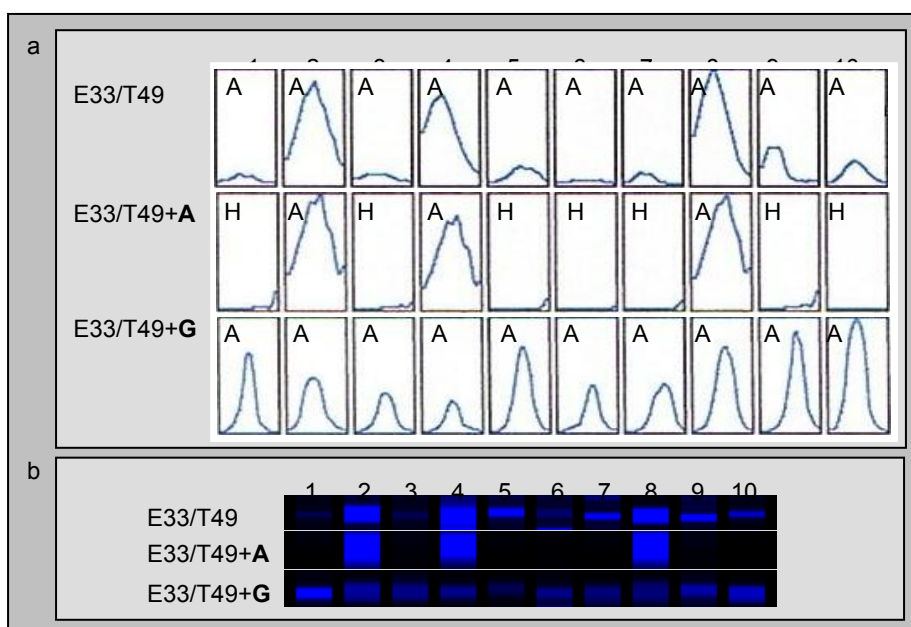


Figura III.9. Patrones de amplificación AFLP de un fragmento de 368 pb en 10 individuos de la familia H8, a partir de la pareja de cebadores E33/T49 y de las parejas E33/T49+C y E33/T49+A. (a) Perfil de picos AFLP resultantes de cada una de las parejas de cebadores. Los códigos A y H indican el fenotipo de presencia o ausencia de banda, respectivamente. (b) Patrón de bandas correspondiente a los picos de la figura a. Las imágenes proceden del software Genographer (Benham J.J. Montana State University 2001).

El efecto de la suma de fragmentos de igual tamaño procedentes de distintas regiones genómicas se puede evidenciar también en algunas bandas monomórficas, en función de las distintas intensidades de las amplificaciones. Bandas de igual tamaño a las bandas monomórficas generadas por la pareja E33/T49, se pueden identificar en los perfiles de las parejas de cebadores E33/T49+N. En algunos casos las bandas monomórficas E33/T49 se correspondieron con bandas E33/T49+N también monomórficas. En otras

ocasiones, se correspondieron en tamaño tanto con bandas monomórficas como con bandas polimórficas generadas por los cebadores E33/T49+N. En estos casos, donde una banda monomórfica resulta de la suma de bandas polimórficas y monomórficas, se pueden apreciar diferencias de intensidad entre los fragmentos amplificados, dependiendo del fenotipo de la banda polimórfica. Un ejemplo es la banda E33/T49.278, donde el efecto de la suma de fragmentos procedentes de distintas regiones (E33/T49+A.278 y E33/T49+G.278) se hace evidente en las distintas intensidades de las bandas (Figura III.10b).



Puesto que las bandas monomórficas no son informativas en un análisis de segregación, estas bandas no introducen ningún error en el proceso de genotipado, sin embargo sirven para constatar la existencia de homoplasias

III. 3. CARTOGRAFÍA GENÉTICA DEL RODABALLO

III.3.1. Análisis dos puntos. Asignación de los marcadores a los grupos de ligamiento

Las relaciones de ligamiento de los 136 marcadores segregantes (112 AFLPs y 24 microsatélite) genotipados en la progenie ginogenética haploide de una hembra de *Scophthalmus maximus* se analizaron con el software MAPMAKER/EXP versión 3.1 (Lander *et al.*, 1987). Los primeros datos de ligamiento se obtuvieron de un análisis dos-puntos con una puntuación mínima de LOD = 3,0 y una frecuencia máxima de recombinación $r = 0,30$. Bajo estas condiciones, se establecieron 22 grupos de ligamiento formados por 97 marcadores (71,3%), mientras que 39 marcadores resultaron no ligados. Incrementando el LOD a 4,0 el marco global de grupos se mantuvo, si bien algunos marcadores (15) ligados a LOD 3,0 perdieron su posición en el mapa. Relajando el criterio de distancia ($r = 0,40$) y manteniendo el LOD 3,0, uno de los marcadores aislados se unió a un grupo de ligamiento ya establecido (E40/T50.172 en el grupo 7).

Puntuaciones LOD por debajo de 3,0 pueden producir ligamientos espúreos, resultantes de la aglomeración de los grupos de ligamiento más grandes con grupos más pequeños. Sin embargo, en algunas ocasiones fue necesario relajar la puntuación mínima de LOD a 2,0 para llegar a un consenso entre el número de datos disponibles y el número de cromosomas de la especie. Esta mayor permisividad en los criterios para establecer ligamiento entre marcadores permitió identificar tres nuevos grupos de ligamiento: grupos 11, 12 y 13, y detectar la asociación de 13 marcadores adicionales a los grupos previamente establecidos.

Finalmente, el mapa genético resultó formado por 25 grupos de ligamiento: 7 grupos de ligamiento mayores (con 5 o más marcadores) y 18 grupos menores (2-4 marcadores). Dicho mapa incluye 118 marcadores ligados (86,7%), mientras que 18 marcadores resultan no ligados (Figura III.11).

Los grupos del mapa se designan con las siglas LG (*Linkage Group*) seguidas de un número en orden decreciente según el tamaño relativo en distancia genética para cada grupo.

III.3.2 Análisis multipunto. Orden de marcadores y distancias de mapa.

Los grupos de ligamiento identificados a valores umbral de $LOD = 3,0$ y $r = 0,40$ fueron ordenados siguiendo un análisis multi-punto. Para este análisis se excluyeron en un principio los marcadores con distorsión de segregación ($0,01 < P < 0,05$), que sólo se posicionaron en el mapa una vez establecido el orden de los marcadores no distorsionados. En cada grupo de ligamiento se obtuvo un mapa de referencia o mapa *framework* en el cual, un determinado orden fue aceptado para un marcador dado, sólo cuando el logaritmo de verosimilitud de ese orden fue al menos 2 veces mayor que el del segundo mejor orden. Los marcadores ordenados en este mapa constituyen los loci *framework*.

Una vez caracterizadas las relaciones de ligamiento entre los marcadores ligados con un $LOD \geq 3,0$, se procedió a ordenar los marcadores ligados con puntuaciones LOD inferiores. Cuando la posición de dichos marcadores pudo ser identificada con un $LOD \geq 2,0$, los nuevos loci se incluyeron en el mapa *framework* (Figura III.11). En todos los casos, estos nuevos loci ocupan posiciones periféricas en los grupos de ligamiento a los que pertenecen. Finalmente, el resto de marcadores ligados al mapa, cuya posición no pudo determinarse con precisión ($LOD < 2,0$) fueron considerados marcadores accesorios y se incluyen entre corchetes al lado del marcador

framework al que se ligán con un mayor LOD, seguidos de la distancia en cM a dicho marcador (Figura III.11).

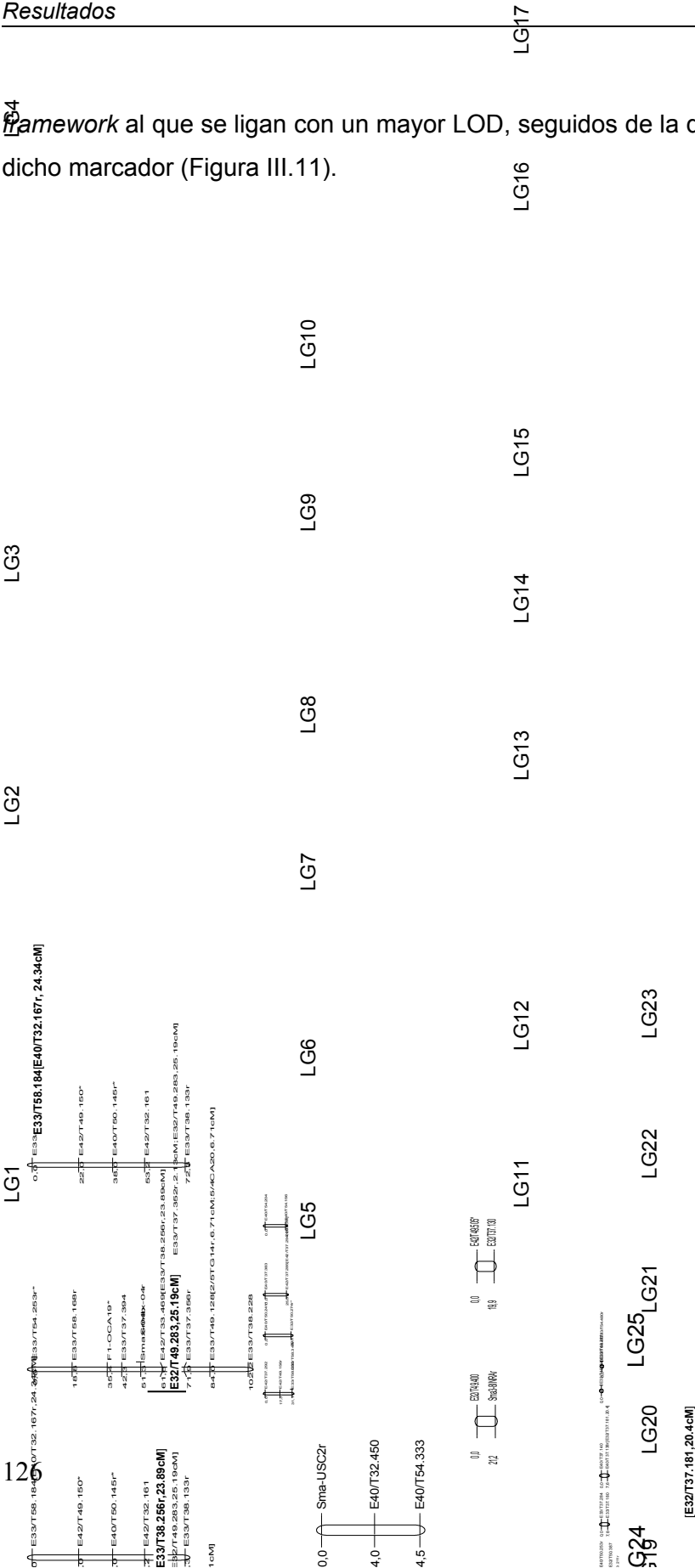
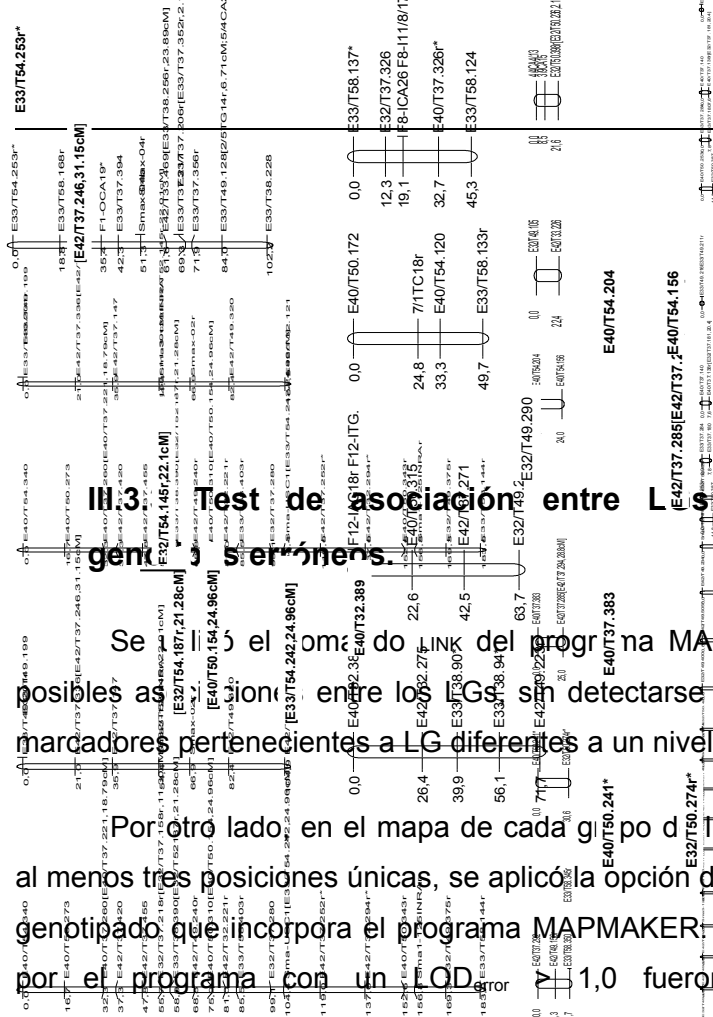


Figura III.11. Mapa de ligamiento de *Scophthalmus maximus*. Los marcadores se representa en el lado derecho de cada LG, mientras que los números a la izquierda indican la distancia genética de cada intervalo (cM Kosambi). Los marcadores recodificados se indican con una *r*. Los loci con una ratio de segregación distorsionada ($0,01 < P < 0,05$), se identifican con un asterisco (*). Los marcadores que se unieron al mapa *framework* con puntuaciones LOD inferiores a 3,0 se representan en letra negra. Los marcadores accesorios se incluyen entre corchetes al lado del marcador *framework* al que se unen con una mayor puntuación LOD. Las líneas verticales en los LG1 y 5 indican los intervalos con una concentración de marcadores superior a la esperada según Poisson ($P > 0,05$).



Por otro lado, en el mapa de cada grupo de ligamiento, constituido por al menos tres posiciones únicas, se aplicó la opción de detección de errores de genotipado que incorpora el programa MAPMAKER. Los genotipos puntuados por el programa con un QD $\geq 1,0$ fueron comprobados en los correspondientes geles a fin de identificar posibles errores en la interpretación de la presencia o ausencia de banda en el caso de marcadores AFLPs, o de presencia de un determinado alelo en el caso de los microsatélites. De esta forma, tras la corrección de errores de genotipado, la longitud total del mapa se redujo en un 4%.

III.3.4. Características del mapa de ligamiento genético de *S. maximus*

Ciento dieciocho marcadores (86,7%; 98 AFLP y 20 microsatélites), distribuidos en 94 posiciones únicas o *bins*, constituyen el mapa genético femenino de *S. maximus*, cubriendo una longitud genética total de 1010,4 cM, según la función de mapa Kosambi (los marcadores accesorios no se incluyeron en el cálculo de la distancia total). Si no se considera la contribución

de los grupos representados por menos de tres *bins*, la longitud del mapa disminuye a 790,6 cM.

Entre los 98 marcadores AFLP mapeados se identificaron 5 pares de alelos de marcadores AFLP codominantes (LG2: E33/T49.199-E33/T49.201r; LG3: E33/T37.206r-E33/T37.217; LG10: E33/T58.345r-E33/T58.350; LG24: E33/T49.211r-E33/T49.216; LG25: E33/T54.477-E33/T54.480r) que constituyen el 4,6% del total de bandas AFLP polimórficas incluidas en el análisis de ligamiento. Dichos marcadores se caracterizan por ser amplificados por la misma combinación de cebadores, diferenciarse en pocos pares de bases a nivel de tamaño y cosegregar en fase de repulsión (Paglia *et al.*, 1998). De esta forma, los grupos 24 y 25 constituidos por pares de marcadores codominantes, no representan verdaderos grupos de ligamiento, sino un par de marcadores no ligados (E33/T49.211r-216, E33/T54.477-480r). Así, el número de marcadores totales del mapa se reduce de 118 a 111.

Los 111 marcadores únicos se distribuyen entonces en 23 grupos de ligamiento que varían en tamaño desde 7,6 cM (LG23) a 183,0 cM (LG1) siendo el tamaño medio de los grupos 43,9 cM (d.e = 41,3). Estos grupos incluyen 92 *bins* con la siguiente distribución de marcadores: 76 *bins* (82,6%) ocupados por un único marcador; 13 *bins* (14,1%) ocupados por dos marcadores; y 3 *bins* (3,2%) por 3 marcadores. Así, el número máximo de marcadores por posición de mapa o *bin* es 3, sin encontrarse agrupaciones de AFLPs como las descritas por otros autores en mapas genéticos basados en este tipo de marcador (Bert *et al.* 1999). El número de marcadores por grupo de ligamiento varía desde 2 (LG22) a 25 (LG1), con una media de 5,1 marcadores por grupo y una distancia genética media entre marcadores adyacentes de 14,6 cM. El intervalo más largo se detectó en el LG11 (30,6 cM), un doblete constituido por marcadores con distorsión de segregación que se unen con una puntuación LOD = 2,05. Sin embargo, en la mayoría de los grupos, el espaciado entre marcadores consecutivos fue menor de 25 cM. Se

exceptúa el LG5 que presentó un largo intervalo sin marcadores o *gap* de 26,4 cM (LOD = 2,73 entre los marcadores E40/T32.389 y E42/T32.275) y el LG12, con un *gap* de 25,0 cM (LOD =2,59). Las Tablas III.9 y III.10 resumen las características del mapa.

Tabla III.9. Características de los LGs del mapa genético femenino de *S. maximus*: Longitud (cM Kosambi), número de marcadores, espaciado medio e intervalo más largo de los LG constituidos por tres o más marcadores.

LG	Longitud (cM)	Nº marcadores framework + accesorios	Espaciado medio (cM)	Intervalo mayor (cM)
LG1	183,0	19 + 6	10,2	18,0
LG2	104,9	7 + 1	17,4	22,5
LG3	102,2	11 + 5	11,3	18,9
LG4	72,3	5 + 1	18,1	22,0
LG5	71,7	5 + 0	17,9	26,4
LG6	63,7	5 + 0	21,2	22,6
LG7	49,7	4 + 0	16,5	24,9
LG8	45,3	6 + 0	11,3	13,6
LG9	44,5	3 + 0	22,2	24,0
LG10	31,7	3 + 0	15,8	17,3
LG15	21,6	3 + 1	10,8	13,1

Tabla III.10. Características globales del mapa de ligamiento femenino de *S. maximus*.

Longitud total del genoma mapeado (cM)	1010,4
Distancia media entre dos marcadores (cM)	14,6
Longitud media de los grupos de ligamiento (cM)	43,9
Nº medio de marcadores por grupo de ligamiento:	
Marcadores <i>framework</i> más pares de marcadores ligados.	4,0
Marcadores <i>framework</i> , pares de marcadores ligados y marcadores accesorios.	4,8

III.3.5. Presencia en el mapa de marcadores con distorsión de segregación

En el presente trabajo, las proporciones de marcadores ligados al mapa total (*framework* más accesorios y pares de marcadores ligados) entre aquellos que se ajustan a las proporciones mendelianas de segregación y los distorsionados fueron muy similares: 86,7% y 87,5%, respectivamente. Sin embargo, la diferencia se hace mayor cuando se compara el porcentaje de

marcadores distorsionados y no distorsionados que se ligan al mapa *framework* ($\text{LOD} \geq 3,0$): 73,5% y 60%, respectivamente. La Tabla III.11 resume la presencia de ambos tipos de marcadores en el mapa. Por otro lado, en cuatro de los siete grupos de ligamiento en los que se distribuyen los marcadores con distorsión de segregación, pares de estos marcadores aparecen en posiciones adyacentes (Figura III.11, LG1: E42/T37.252r*-E42/T32.294r*; LG4: E42/T49.150*-E40/T50.145r*; LG5: E33/T38.90*-E33/T38.94*; LG11: E40/T50.241*-E32/T50.274r*).

Tabla III.11. Presencia en la mapa de marcadores con segregación alélica mendeliana y segregación distorsionada. Se indica el número de marcadores que mostraron ajuste ($P > 0,05$) o desviación ($0,01 < P < 0,05$) respecto a la ratio de segregación mendeliana entre los marcadores *framework*, accesorios, pares de marcadores ligados y marcadores no ligados del mapa genético femenino de *S.maximus*.

Características	Nº marcadores Segr. Mendel. + No Segr. Mendel..
Marcadores <i>framework</i>	
AFLP	51 + 9
Microsatélite	15 + 1
Marcadores accesorios	
AFLP	14 + 0
Microsatélite	2 + 0
Dobletes	
AFLP	14 + 3
Microsatélite	3 + 0
Marcadores no ligados	
AFLP	12 + 2
Microsatélite	4 + 0

III.4. DISTRIBUCIÓN DE LOS MARCADORES A LO LARGO DEL MAPA.

III.4.1. Distribución de marcadores dentro de los grupos de ligamiento.

Cuando en el análisis se consideraron exclusivamente los marcadores *framework*, no se detectó ningún cluster de marcadores dentro de los grupos de ligamiento y la frecuencia observada de cada tipo de intervalo se ajustó a lo esperado según la distribución de Poisson a un nivel de significación $P < 0,05$. Sin embargo, cuando se incluyeron en el análisis los marcadores accesorios,

se detectaron en los grupos de ligamiento LG1 y LG3 dos intervalos con un exceso de marcadores según lo esperado en una distribución aleatoria de polimorfismos. En ambos LGs, dichos agrupamientos resultaron constituidos exclusivamente por marcadores AFLP (Figura III.11).

III.4.2. Distribución de marcadores entre los grupos de ligamiento.

III.4.2.1. Distribución de marcadores independientemente de su origen y composición.

El coeficiente de correlación de Pearson (R) entre el número de marcadores por grupo de ligamiento y el tamaño de los grupos (cM) reveló una correlación significativa a un nivel de significación del 1% (Figura III.12). R tomó un valor de 0,95 al incluir en el análisis sólo los marcadores *framework*. Cuando se consideraron todos los marcadores ligados en el mapa, incluyendo los marcadores accesorios, disminuyó ligeramente ($R = 0,92$).

El test χ^2 identificó diferencias significativas ($P < 0,01$) entre las frecuencias observadas y esperadas de intervalos de 10 cM con 1 y 5 marcadores en los mapas *framework* y total. En ambos casos, la frecuencia de dichos intervalos fue mayor de la esperada según la distribución de Poisson.

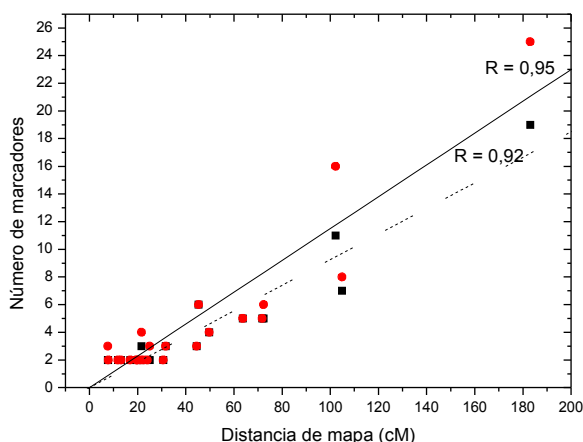


Figura III.12. Correlación entre el número de marcadores y el tamaño de los grupos de ligamiento (LGs). Los puntos negros y la línea de regresión continua representan la relación entre el número de marcadores *framework* y el tamaño de los LGs ($R = 0,95$). Los puntos rojos y la línea discontinua representa la relación entre el número de marcadores totales y el tamaño de los LGs ($R = 0,92$).

III.4.2.2. Distribución de marcadores en función de su origen y composición nucleotídica.

Marcadores Microsatélite

La distribución de los loci microsatélite en función de la complejidad del motivo de repetición, se evaluó a lo largo de los grupos de ligamiento con más de 3 posiciones únicas o *bins* aplicando un test χ^2 de contingencia, sin encontrarse una asociación significativa ($P > 0,05$) entre los marcadores con estructuras comunes en cuanto al motivo de repetición.

Los mismos marcadores microsatélites fueron también clasificados en función de su origen (1, Coughlan *et al.*, 1996; 2, Estoup *et al.*, 1998; 3, Iyengar *et al.*, 2000; 4, Pardo *et al.*, 2005) y se analizó la aleatoriedad de la distribución de marcadores aislados a partir de las distintas genotecas. También en este caso, el análisis estadístico apoya una la distribución aleatoria de los loci microsatélite analizados ($P > 0,05$).

Marcadores AFLP

Mediante el test χ^2 de contingencia se identificaron diversos loci AFLP amplificados a partir de una misma pareja de cebadores cuya concentración a lo largo de los grupos de ligamiento difirió significativamente ($P < 0,05$) de la esperada por azar. Fue éste el caso de los marcadores amplificados a partir de las parejas E32/T50, E33/T37, E33/T38, E33/T58 y E42/T37 que mapean en los grupos de ligamiento LG15, LG3, LG5, LG8 y LG2, respectivamente (Figura III.11, Figura III.13).

Tanto en el mapa *framework* como en el mapa total (considerando marcadores accesorios y marcadores *framework*), la concentración de estos marcadores en dichos grupos superó la esperada bajo la suposición de aleatoriedad de la distribución.

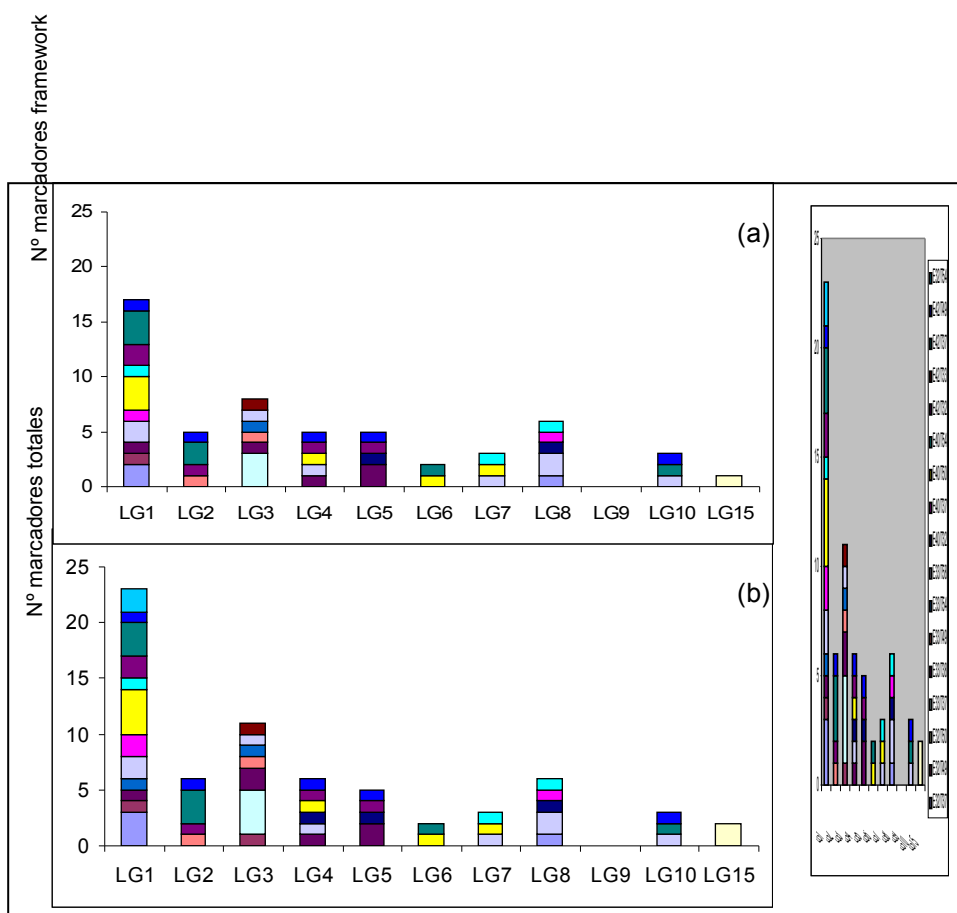


Figura III.12. Distribución de marcadores AFLP a lo largo de los grupos de ligamiento del mapa genético de *S. maximus*. (a) Marcadores framework. (b) Marcadores accesorios y framework. Los marcadores amplificados por una misma combinación de cebadores se representan con un código de color, que se indica en el lado derecho del gráfico.

La longitud genómica observada (G_o) del mapa de ligamiento genético de *S. maximus* comprende 790,6 cM, cuando se consideran únicamente los grupos de ligamiento *framework* y alcanza los 1010,4 cM, cuando se consideran en este cálculo los pares de marcadores ligados en dobletes (Tabla III.12). Estas longitudes resultan de la suma de las distancias de todos los intervalos mapeados. Sin embargo, estas distancias derivadas de un mapa

de ligamiento parcial no representan la longitud total del mapa de rodaballo. Es necesario recurrir a estimadores que a partir de los datos parciales faciliten una predicción aproximada de la longitud del genoma de esta especie.

La longitud genómica (Ge) de *S. maximus* fue estimada, en primer lugar, según el método de Chakravarti *et al.* (1991). Considerando sólo los marcadores *framework*, la longitud esperada del genoma de rodaballo sería de 1231,1 cM (Tabla III.12). El valor Ge estimado mediante el método de Fishman *et al.* (2001), que incluye en el cálculo de la longitud la distancia teórica desde el último marcador de cada cromosoma hasta su telómero, fue de 1065,9 cM (Tabla III.12). A partir de las dos estimas se obtiene un valor promedio de longitud genómica esperada en esta especie de 1148,5 cM (Tabla III.12).

La fracción de la longitud genómica esperada Ge que cubre la longitud observada Go indica la cobertura genómica del mapa de ligamiento (Co). Esta cobertura puede calcularse directamente como la *ratio* entre ambos valores de longitud genómica (Go/Ge). Cuando en el cálculo de Go se incluyen sólo los marcadores *framework* (Go_f) la cobertura genómica del mapa (Co_f) es del 68,8%. Si se considera en el cálculo la distancia total de mapa incluyendo los marcadores *framework* y grupos formados por pares de marcadores ligados (Go_t), la cobertura del mapa total (Co_t) aumenta hasta el 87,9% (Tabla III.12).

Tabla III.12. Longitud del genoma y cobertura del mapa genético femenino de *S. maximus*

Longitud genómica	<i>S. maximus</i>
Longitud genómica observada (cM) ^a	
Go_f	790,6
Go_t	1010,4
Longitud genómica esperada (cM) ^b	
Ge_c	1231,1
Ge_F	1065,9
Ge	1148,5
Cobertura genómica (%) ^c	
Co_f	68,8%
Co_t	87,9%
Ce_L	69,0%

^a Go_f : longitud genómica observada a partir del mapa *framework*; Go_t : longitud genómica observada total (mapa *framework* más dobletes).

^b Ge_c : longitud genómica esperada según Chakravarti *et al.* (1991); Ge_f : longitud genómica esperada según Fishman *et al.* (2001); Ge : longitud genómica esperada promedio.

^c Co_f : cobertura genómica del mapa *framework*; Co_t : cobertura genómica observada del mapa total (*framework* más dobletes); Ce_L : cobertura genómica teórica según Lange y Boehnke (1982).

A partir del número de marcadores mapeados en el mapa *framework* y de los datos de distancias entre dichos marcadores y el número de cromosomas haploides de la especie en cuestión, es posible hacer una estima de la cobertura teórica que se podría esperar de un determinado mapa de ligamiento parcial (Tabla III.12). Considerando el método propuesto por Lange y Boehnke (1982) la cobertura genómica esperada (Ce_L) a partir del mapa de rodaballo sería del 69%. La cobertura teórica resultó muy próxima a la cobertura obtenida cuando se considera la longitud genómica observada (Co_f) del mapa *framework*. (Tabla III.12).

IV. DISCUSIÓN

La discusión de este estudio se ha estructurado en tres apartados que analizan las principales aportaciones de este trabajo de Tesis: (1) la estrategia de mapeo a partir de una población de embriones ginogenéticos haploides; (2) el desarrollo de nuevos marcadores moleculares en rodaballo (*Scophthalmus maximus*); y (3) la presentación del primer mapa de ligamiento de esta especie.

IV.1. ESTRATEGIA DE MAPEO: FAMILIAS DE GINOGENÉTICOS HAPLOIDES

Desde que en 1994, Lie y colaboradores demostraron las ventajas del uso de familias ginogenéticas haploides en el análisis de ligamiento en peces (particularmente en el salmón atlántico (*Salmo salar*)), la misma estrategia se ha empleado con éxito para el desarrollo de mapas genéticos de moderada densidad en otras especies piscícolas como el pez cebra (*Danio rerio*, Postlethwait *et al.*, 1994; Johnson *et al.*, 1996), el salmón atlántico (*Salmo salar*, Slettan *et al.*, 1997), la tilapia (*Oreochromis niloticus*, Kocher *et al.* 1998), la carpa (*Cyprinus carpio*, Sun y Liang, 2004) o el pez gato (*Ictalurus punctatus*, Poompuang y Na-Nakorn, 2004).

Cuando se plantea un estudio de ligamiento genético, la utilización de familias segregantes haploides frente a una población de referencia convencional presenta una serie de ventajas: i) ahorro de tiempo y costes, ya que se evita la construcción y mantenimiento de pedigrís informativos; ii) posibilidad de obtener una mayor precisión en la estima de la tasa de recombinación gracias a la disponibilidad de amplios tamaños muestrales; y iii) simplicidad del genotipado del material haploide cuando se trabaja con marcadores dominantes como RAPDs (*Random Amplified Polymorphic DNA*) o AFLPs (*Amplified Fragment Length Polymorphism*).

En el presente estudio, se decidió seguir esta estrategia de mapeo considerando la problemática concreta del rodaballo. En primer lugar, la falta de líneas puras divergentes o pedigrís seleccionados, que permitieran la

obtención de familias de referencia apropiadas para el análisis de ligamiento, frente a la disponibilidad de la tecnología necesaria para la inducción de ginogénesis en esta especie (Castro *et al.*, 2003; Piferrer *et al.*, 2004). Por otro lado, hasta muy recientemente, el número de marcadores publicados en esta especie era muy escaso, con apenas una treintena de marcadores microsatélites descritos en la literatura (Coughlan *et al.*, 1996; Estoup *et al.*, 1998; Iyengar *et al.*, 2000; Bouza *et al.*, 2002). Estas circunstancias motivaron la aplicación de la tecnología AFLP, que resulta muy eficaz en cuanto al número de polimorfismos obtenidos en relación al tiempo y coste económico empleados, si bien, los marcadores que se obtienen segregan fundamentalmente de forma dominante. El seguimiento de la transmisión de este tipo de marcadores se simplifica mucho cuando se trabaja con familias haploides. En el caso de familias ginogenéticas, la segregación alélica de una madre heterocigótica puede determinarse directamente en su descendencia ginogenética haploide.

Sin embargo, no todo son ventajas en el uso de este tipo de material. En primer lugar, hay que considerar que las estimas obtenidas de la tasa de recombinación son exclusivamente estimas femeninas y que, en la mayoría de los Vertebrados estudiados, incluidos los peces, las tasas de recombinación difieren entre sexos (Dib *et al.*, 1996; Dietrich *et al.*, 1996; Mikawa *et al.*, 1999; Sakamoto *et al.*, 2000; Singer *et al.*, 2002). Es por tanto importante considerar que el mapa obtenido en este estudio es un mapa genético femenino del rodaballo, que puede presentar variaciones en las frecuencias de recombinación respecto al mapa masculino de la misma especie. Estas diferencias se discutirán más ampliamente en apartados sucesivos.

En segundo lugar y no menos importante, uno de las mayores limitaciones de estos mapas contruidos a partir de embriones es la imposibilidad de mapear caracteres fenotípicos del individuo adulto. El síndrome haploide impide un desarrollo normal de los embriones y estos mueren antes de llegar a eclosionar. Así, la información de los marcadores se

limita casi exclusivamente a polimorfismos de ADN amplificables vía PCR, pero no proporcionan información acerca de la localización en el mapa de genes que codifiquen alozimas, color o sexo del individuo. Por otro lado, si se pudiera medir algún rasgo fenotípico del embrión, sólo se estimaría el efecto materno de los genes subyacentes o QTL (*Quantitative Trait Loci*) en el caso de los caracteres complejos. Dado el interés productivo de muchos rasgos cuantitativos, en una segunda generación de mapas, sería interesante disponer de familias ginogenéticas de dobles haploides a fin de obtener paneles diploides que pueden llegar a desarrollarse como individuos adultos y manifestar caracteres fenotípicos medibles.

Otro problema relacionado con el genotipado de material haploide es la cantidad de ADN que puede obtenerse de cada embrión. Lie y colaboradores (1994) destacaban las ventajas del uso de embriones haploides frente al genotipado de células espermáticas individuales en relación a la mayor cantidad de ADN que puede extraerse de los primeros. Sin embargo, trabajos más recientes insisten en la escasa cantidad de ADN que puede obtenerse de las progenies haploides, especialmente problemática cuando deben genotiparse en un mismo individuo cientos de marcadores, que en ciertos casos, como los de RAPDs y AFLPs, presentan requerimientos de cantidad y calidad de ADN de partida relativamente elevados (Chagnè *et al.*, 2002; Sun y Liang, 2004). En el presente estudio, se tuvo en cuenta esta dificultad al aplicar el protocolo de extracción de ADN basado en la sucesión de etapas de lavado con fenol: cloroformo: alcohol isoamílico. La optimización de dicho protocolo permitió mejorar el rendimiento de la extracción en un 40%, sin embargo, la cantidad de ADN obtenida continuaba siendo crítica para la caracterización del número de marcadores necesario para el desarrollo del mapa. El problema de la escasez de ADN se solventó finalmente con la aplicación del protocolo de amplificación genómica GenomiPhi™, que a partir de 10 ng de ADN de cada embrión permitió disponer de un reservorio de ADN del orden de microgramos

para la caracterización de nuevos marcadores y para una posible ampliación de este mapa en el futuro.

Finalmente, algunos autores han manifestado cierta cautela en relación al efecto que los genes deletéreos podrían ocasionar en las poblaciones de mapeo haploides. En dichas poblaciones, los alelos recesivos se manifiestan fenotípicamente, pudiendo ocasionar distorsión de segregación en caso de tratarse de alelos letales. Lie *et al.* (1994) demostraron mediante una simulación computacional que la presencia de un gen letal podría llevar a una distorsión de la segregación en los marcadores individuales y a la eliminación de algún haplotipo, pero en ningún caso influiría en la estima de la tasa de recombinación. Posteriormente, algunos autores han descrito casos más o menos serios de distorsión de segregación en marcadores genotipados en familias haploides (Kocher *et al.*, 1998; Poompuang y Na-Nakorn, 2004). Sin embargo, estos mismos investigadores justifican el origen de esta distorsión en problemas derivados de errores de genotipado más que en la presencia de alelos deletéreos.

En nuestro estudio, la notable diferencia entre marcadores AFLP y microsatélite en cuanto a la proporción de marcadores con segregación no mendeliana (25,7% y 4,1 %, respectivamente) así como, una distribución no agrupada de la mayoría de los marcadores distorsionados a lo largo del mapa de ligamiento, indican posibles errores de genotipado en la caracterización de los marcadores AFLP, más que la presencia de alelos deletéreos recesivos, como causa más probable de las distorsiones observadas.

IV.2. HERRAMIENTAS MOLECULARES: MARCADORES MICROSATÉLITE Y AFLP

Dado el amplio espectro de marcadores disponibles actualmente, Liu y Cordes, en su revisión del 2004 sobre la aplicación de los marcadores moleculares en la acuicultura, destacaban la necesidad de conocer, antes de

empezar una investigación genómica, cual es el tipo de marcador que mejor se adapta a los objetivos específicos del estudio considerando, en primer lugar, la información genética de la que se dispone para la especie objeto de estudio.

En el caso concreto de rodaballo, se utilizaron marcadores microsatélite y AFLP para la construcción de un mapa genético de moderada densidad. Los primeros son los marcadores de elección en cualquier proyecto de mapeo en el que se disponga *a priori* de información sobre las secuencias de ADN. En el caso de rodaballo, la información genómica era escasa y fue necesario recurrir a la amplificación cruzada y al desarrollo de una librería genética enriquecida en secuencias repetidas en tándem (SSR: *Simple Sequence Repeat*) para el desarrollo de nuevos loci microsatélite. Estas premisas hacen de la tecnología microsatélite, basada en la secuenciación y el diseño de cebadores específicos de locus, un procedimiento costoso. Los AFLP surgen entonces como una alternativa a los marcadores específicos de secuencia. Estos marcadores no necesitan información acerca la secuencia de ADN y por tanto, pueden aplicarse inicialmente a cualquier especie en estudio, se obtienen a partir de un protocolo relativamente sencillo y tienen una elevada razón múltiplex (número de marcadores obtenidos por combinación de cebadores). En un proyecto de mapeo, en el que el fin más inmediato es obtener un mapa detallado en el que poder detectar ligamiento entre marcadores y loci implicados en caracteres cuantitativos (QTL), lo ideal es combinar ambos tipos de marcadores (Nichols *et al.*, 2003; Watanabe *et al.* 2004). Los microsatélites dan robustez al mapa y posibilitan futuros mapeos comparativos. Además permiten el seguimiento de la naturaleza multialélica de los QTL (Sewell *et al.*, 2000), lo cual no sería posible con marcadores dialélicos dominantes como los AFLP. Por su parte, estos últimos posibilitan con un coste moderado, la construcción de mapas genéticos relativamente densos, facilitando el análisis de asociación entre marcadores y la estimación del tamaño genómico en especies poco conocidas a nivel de su ADN (Naruse *et al.*, 2000; Cervera *et al.*, 2001; Liu *et al.*, 2003)

IV.2.1. Marcadores microsatélite en *Scophthalmus maximus*

A continuación se discuten las ventajas e inconvenientes de las estrategias que se siguieron en este trabajo para la identificación de nuevos loci microsatélite en rodaballo: amplificación cruzada y desarrollo de una librería parcialmente enriquecida en secuencias SSR.

Amplificación Cruzada

La amplificación cruzada es una práctica habitual cuando se trabaja con loci microsatélite, ya que son loci abundantes en todas las especies y han demostrado estar bien conservadas entre linajes próximos (Rico *et al.*, 1996; Angers y Bernatchez, 1997; Rajapakse *et al.* 2001; Pinera *et al.*, 2006). En el caso de la familia Scophthalmidae, a la que pertenece el rodaballo, no se disponía de otras especies en las que se hubieran caracterizado marcadores microsatélite y los cebadores testados debieron seleccionarse entre dos especies pertenecientes a una familia filogenéticamente próxima, Pleuronectidae: *Pleuronectes platessa* (platija japonesa, Hoarau *et al.*, 2002) e *Hippoglossus hippoglossus* (fletán, Coughlan *et al.*, 2000). Estudios previos, tanto en *P. platessa* como en *H. hippoglossus* demostraron amplificaciones exitosas a partir de cebadores diseñados para rodaballo en un 22,2% de los loci testados (Iyengar *et al.*, 2000). En el presente estudio, sólo un 11% de los cebadores amplificaron secuencias en *S. maximus*, sin embargo, las amplificaciones fueron inespecíficas y de escasa reproducibilidad. El poco éxito de las amplificaciones cruzadas podría indicar una escasa conservación de las secuencias flanqueantes de los loci microsatélite entre las dos familias de peces planos. Estos resultados apoyan las observaciones de Bouza *et al.* (2002) y Castro *et al.* (2006), que sugieren una baja conservación de los loci microsatélite entre las especies de este grupo de peces.

Librerías Genómicas Enriquecidas en STRs

Una vez descartada la amplificación cruzada para identificar marcadores microsatélite en rodaballo, el desarrollo de una librería genómica a

partir de un individuo de esta especie se presentó como la estrategia más adecuada para la obtención de dichos marcadores. El desarrollo de genotecas enriquecidas en secuencias STR ha demostrado ser, a pesar de su coste, la estrategia más efectiva para el desarrollo de nuevos marcadores microsatélite en una especie dada (Ostrander *et al.*, 1992; Kijas *et al.*, 1994; Pardo *et al.*, 2005). Esta afirmación la constata el creciente número de publicaciones de nuevos marcadores microsatélite en distintas especies en revistas que se dedican en su totalidad (*Molecular Ecology Notes*) o en parte de sus secciones a este tipo de artículo técnico (Animal Genetics, Conservation Genetics, etc.) Una consulta de los artículos publicados en *Molecular Ecology Notes* a lo largo de los últimos 2 años, confirma la difusión y la eficiencia de estas librerías genómicas enriquecidas para la identificación de estos loci (Baranski *et al.*, 2006; Castaño-Sánchez *et al.*, 2006; Ha *et al.*, 2007; revisión en Zane *et al.*, 2002).

Particularmente en rodaballo, para la construcción de la librería parcialmente enriquecida en secuencias STR se utilizó la sonda [CA]₁₇. La repetición dinucleotídica [CA]_n es la más usada en la identificación de estos marcadores, ya que corresponde al motivo microsatélite más abundante en los genomas de Vertebrados (Lee *et al.*, 1996). Concretamente, se calcula que la repetición CA se encuentra cada 5 kb en el genoma de rodaballo (Iyengar *et al.*, 2000). Siguiendo esta estrategia, se identificaron 8 loci microsatélite, de los cuales 5 resultaron polimórficos en una muestra poblacional de rodaballo cantábrico. El rendimiento obtenido fue similar al de otros estudios a partir de las primeras genotecas realizadas en esta misma especie: 4 marcadores microsatélite aislados por Coughlan *et al.* (1996) y 10 aislados por Iyengar *et al.* (2000); y se asemeja al obtenido en otras especies de peces planos: 11 marcadores en *H. hippoglossus* (Coughlan *et al.*, 2000) y 7 marcadores en *P. platessa* (Hoarau *et al.*, 2002).

El número medio de alelos para los loci microsatélite aislados a partir de la presente genoteca de rodaballo (6,4) es similar al obtenido por Coughlan *et al.* (1996) e Iyengar *et al.* (2000) para la misma especie (6,7 y 6,5, respectivamente). Sin embargo, el rango del número de alelos por locus (de 4 a 10) descrito en estos trabajos, fue menor al detectado en el presente estudio (de 2 a 15). La escasa variabilidad de algunos de los nuevos loci microsatélite se reflejó también en la genotipado de la hembra H8, que resultó heterocigótica sólo para los marcadores más polimórficos, los loci *Sma*-USC1 y *Sma*-USC2. En el locus *Sma*-USC2 se detectó un déficit de heterocigotos en el análisis poblacional, lo que podría deberse a la presencia de alelos nulos. En el caso de la familia de referencia, la presencia de alelos nulos quedó descartada al ser la hembra H8 heterocigótica para dicho locus.

Los estudios de Brooker *et al.* (1994) mostraron que los microsatélites de teleosteos son generalmente más largos que los de cualquier otro genoma de Vertebrados. Esta observación se confirmó en trabajos posteriores, que constataron la presencia en peces de repeticiones $[CA]_n$ generalmente mayores que las de los genomas de mamíferos (Weber, 1990; Wintero *et al.*, 1992; Rico *et al.*, 1997). En el caso de los nuevos microsatélites identificados en rodaballo, no se observaron repeticiones $[CA]_n$ particularmente largas, si bien, cabe destacar la notable amplitud del rango alélico de dos de los marcadores caracterizados, el *Sma*-USC1 y *Sma*-USC2, que presentaron una variación de tamaño de 38 y 40 pb, respectivamente. En la bibliografía de peces, se ha informado con anterioridad de la presencia de grandes diferencias en el tamaño de los alelos microsatélite (Liu y Cordes, 2004). Concretamente en peces planos, se han descrito rangos alélicos de más de 100 pb de diferencia para *S. maximus* (Estoup *et al.*, 1998), *P. olivaceus* (Sekino y Hara, 2000) y *P. platessa* (Hoarau *et al.*, 2002). En el caso de *H. hippoglossus* las diferencias fueron más discretas, si bien en prácticamente la mitad de los loci descritos, la diferencia en el rango alélico superó las 40 pb (Coughlan *et al.*, 2000). Es precisamente esta característica de los loci

microsatélite para admitir cambios en el número de las unidades de repetición, lo que los convierte en marcadores con un gran número de alelos por locus y por tanto, con un alto PIC (*Polymorphic Information Index*), el mayor en comparación con cualquier otro marcador de ADN (Liu y Cordes, 2004).

IV.2.2. Marcadores AFLP en *Scophthalmus maximus*

Las características más ventajosas de la metodología AFLP (Vos *et al.*, 1995) permiten superar algunas de las principales dificultades relacionadas con la aplicación de la tecnología microsatélite: no requiere información previa a cerca de la secuencia de ADN; tiene una alta razón múltiplex (Powell *et al.*, 1996; Pejic *et al.*, 1998); su aplicación es rápida (una única persona en unos pocos meses puede obtener del orden de cientos de marcadores AFLP); y tiene un coste por marcador relativamente bajo. Estas características, junto con su alta reproducibilidad, han sido determinantes en la amplia difusión de estos marcadores tanto para la construcción de mapas genéticos en especies poco conocidas a nivel molecular (Haanstra *et al.*, 1999; Vuylsteke *et al.*, 1999), como para aumentar la densidad y cobertura de mapas preexistentes basados en marcadores RFLP o microsatélite (Castiglioni *et al.* 1999; Saal y Wricke, 2002; Nichols *et al.*, 2003).

El alto potencial de la tecnología AFLP para detectar polimorfismos genómicos y su relativamente fácil aplicación a cualquier organismo, llevó a algunos autores a considerar que estos marcadores llegarían a desplazar, al menos en algunas aplicaciones, a otros marcadores de alta resolución tales como los RFLPs y microsatélites (Mueller y Wolfenbarger, 1999). Sin embargo, casi una década después, y a pesar del extendido uso de los marcadores AFLP, no se han cumplido las perspectivas más optimistas y los proyectos de mapeo genético más ambiciosos invierten sus esfuerzos en el desarrollo de marcadores microsatélite y SNPs (Shimoda *et al.*, 1999, Sakamoto *et al.*, 2000; Kai *et al.*, 2005; Lee *et al.*, 2005). La clave que marca la diferencia entre AFLPs y microsatélites está en la naturaleza dominante de los primeros, respecto a la

expresión codominante de los segundos, así como en el mayor nivel de polimorfismo de los microsatélite.

Sin embargo, otros factores, generalmente no reconocidos por los autores, podrían haber influido en limitar la aplicabilidad de los marcadores AFLP. La eficiencia de la metodología AFLP para la identificación de polimorfismos podría verse afectada por el diseño de la familia de referencia, por la elección de la pareja de endonucleasas y por la secuencia de bases selectivas incluidas en los cebadores. El efecto de los tres elementos citados podría controlarse cuidando el diseño del protocolo AFLP. Sin embargo, existen otros factores de más difícil control, que influyen en la reproducibilidad de los perfiles AFLP y que con frecuencia se traducen en errores de genotipado. Algunos implican aspectos técnicos como la calidad y cantidad del ADN de partida, defectos en la migración de fragmentos, amplificaciones inespecíficas u homoplasia de tamaño. Otros serían más subjetivos como el criterio y la experiencia del investigador encargado de la caracterización de los nuevos marcadores.

A continuación se discuten la productividad y reproducibilidad de los perfiles AFLP, así como los factores que pudieron influir en el mayor o menor éxito de la aplicación de esta tecnología a embriones haploides de rodaballo.

Factores implicados en la productividad del protocolo AFLP

Familia de referencia. Como ya se comentó en un apartado previo, la principal ventaja de los paneles de mapeo haploides en relación al uso de marcadores AFLP es la propia dotación cromosómica n de la descendencia, ya que pone al descubierto cualquier polimorfismo AFLP, independientemente del carácter dominante de estos marcadores. Ésta característica de las familias haploides podría explotarse mejor si se dispusiera en rodaballo de líneas endogámicas seleccionadas. El diseño de cruces experimentales, que maximizaran la heterozigosis en los parentales de la familia de referencia se reflejarían en tasas de polimorfismo mayores a las detectadas en la

descendencia de individuos salvajes (Pérez *et al.*, 2004; Watanabe *et al.*, 2004). En el presente estudio, la hembra H8 fue seleccionada entre un conjunto de individuos de una población salvaje de rodaballo atlántico en función de su nivel de heterozigosis para un conjunto de loci microsatélite. Si bien ésta selección previa debería aumentar el nivel de polimorfismo de la familia de referencia frente a otras familias de la población salvaje, la eficiencia de los cebadores AFLP en la detección de polimorfismos seguirá dependiendo de la variabilidad natural de la especie.

Digestión genómica. Enzimas de restricción. Los patrones AFLP obtenidos fueron en general amplios, con una distribución de tamaños de banda entre 50 y 650 pb. Las bandas más abundantes fueron las de menor tamaño, disminuyendo su frecuencia a medida que aumentaba su peso molecular. Esta disminución fue paulatina, hubo una gradación y no aparecieron bandas aisladas de excesivo tamaño e intensidad. Todo ello indica una digestión genómica completa por parte de los enzimas *EcoRI* y *TaqI*. Dicha combinación, que había sido seleccionada siguiendo las recomendaciones de Vos y Kupier (1996) para la digestión enzimática de genomas animales, permitió identificar una media de 7,7 (d.t: 3,1) bandas polimórficas en cada uno de los perfiles AFLP. Este valor se encuentra dentro del rango descrito en otras especies de peces: 9,3 marcadores en tilapia (*Oreochromis niloticus*, Kocher *et al.*, 1998); 11,8 en la trucha arco iris (*Oncorhynchus mykiss*, Young *et al.*, 1998); 5,5 en platija (*Pleuronectes platessa*, Coimbra *et al.*, 2003); 9,4 en el pez gato (*Ictalurus punctatus*, Liu *et al.*, 2003). Entre los trabajos citados, sólo en el de la trucha arco iris se utilizó la combinación enzimática *EcoRI/TaqI*, los demás emplearon la combinación clásica *EcoRI/MseI*. Esta última pareja de enzimas fue la propuesta en el protocolo original por Vos *et al.* (1995), no obstante, actualmente se recomienda reservar su uso a la digestión de genomas vegetales (Vos y Kupier, 1996), ya que en animales produce un excesivo número de fragmentos que complican la interpretación de los perfiles AFLP (Ajmone-Marsan *et al.*, 1997). Para entender la diferencia entre los

perfiles AFLP generados por una y otra combinación, ha de considerarse tanto la secuencia de reconocimiento de las enzimas de corte frecuente *MseI* (TTAA) y *TaqI* (TCGA), como las características genómicas del organismo en cuestión. En el genoma de Vertebrados, rico en secuencias no codificadoras compuestas mayoritariamente por nucleótidos AT, muchos puntos serán susceptibles de corte por parte del enzima *MseI*, generando abundantes fragmentos de pequeño tamaño. El enzima *TaqI* con una diana de corte de composición equilibrada entre bases púricas y pirimidínicas, cortará estos genomas con una frecuencia algo menor, generando fragmentos con una distribución de tamaños más amplia. De esta forma, el uso de *TaqI/EcoRI* supone una mejora en el análisis de los genomas complejos animales en términos de calidad del perfil (a través de la reducción del número de fragmentos) y de polimorfismos detectados (presumiblemente a través del marcaje de dinucleótidos CpG). Su adaptación a un sistema de detección automatizado como el CEQ2000 (Papa *et al.*, 2005), incrementa su sensibilidad y precisión en la detección de los marcadores generados, lo que contribuirá a la difusión de este protocolo AFLP *TaqI/EcoRI* entre las especies animales.

Composición de los cebadores y amplificación de fragmentos. Una vez confirmado el éxito de la etapa de restricción, otros factores van a ser los responsables de las variaciones en la media de bandas totales y polimórficas que se encontraron entre los 17 perfiles AFLP analizados. Entre estos factores, el más evidente es la composición de bases selectivas que diferencia a los cebadores AFLP.

En cuanto al número de bandas totales, los perfiles AFLP más pobres (con menos de 50 bandas: E32/T50, E32/T49, E32/T54, E33/T54, E33/T58, E42/T32, E42/T33, E42/T37) correspondieron a los amplificados por parejas de primers que incluían algún cebador con una presencia preponderante de AT entre sus bases selectivas (E32, E33, E42, T32, T37). En el otro extremo de la distribución, fueron las parejas de cebadores que incluían al primer E40 las que

presentaron un mayor número de bandas totales (media de 63,7 bandas por perfil AFLP). Curiosamente es éste el cebador *Eco* que contiene un mayor número de bases GC en su extensión selectiva (AGC). La tendencia observada de un aumento en el número de bandas amplificadas con el contenido de bases GC en las extensiones selectivas de los cebadores, apoya también la influencia de la composición nucleotídica de los cebadores en su potencial de amplificación. Vos *et al.* (1995) en la presentación de la metodología AFLP hacían referencia al control que se podría ejercer en el número de bandas amplificadas a través de la composición de bases selectivas y advertía sobre la posible reducción del número de bandas AFLP como consecuencia de la utilización de trinucleótidos raros.

En cuanto a la capacidad de detección de polimorfismos, las diferencias demostradas por las distintas combinaciones de cebadores dependerán tanto de su potencial para la amplificación de secuencias (puesto que cuantos más fragmentos amplifiquen más posibilidades tendrán de evidenciar mutaciones en las dianas de restricción), como de la propia capacidad de mutación de las secuencias complementarias a las bases selectivas del cebador. De acuerdo con el primer supuesto, en este estudio se ha evidenciado una correlación positiva entre el número de bandas amplificadas y el de polimorfismos detectados por combinación de cebadores. En el segundo caso, considerando que la presencia de bases GC favorece la identificación de polimorfismos por parte del enzima *TaqI*, cabría esperar el mismo efecto en los cebadores que incluyesen dichos nucleótidos entre sus bases selectivas, si bien, sólo unos pocos estudios en vegetales constatan tal influencia (Keim *et al.*, 1997; Remington *et al.*, 1999). En el presente estudio, no pudo establecerse ninguna relación entre la composición de las extensiones selectivas y una mayor o menor ocurrencia de polimorfismos. Dado que en Vertebrados, las secuencias genómicas más ricas en CpG son las secuencias codificadoras (Nekrutenko y Li, 2000; Belle *et al.*, 2002; Fortes *et al.*, 2007), y por tanto, las regiones más reticentes al cambio, podría considerarse la posibilidad de que esta

circunstancia moderase la tendencia mutacional de las bases CpG. En todo caso, en el presente estudio, la representación de las combinaciones de cebadores con más de tres bases selectivas GC fue demasiado escasa para sustentar tal hipótesis.

Factores que afectan a la reproducibilidad de los perfiles AFLP

Del total de 186 bandas AFLP polimórficas, 74 fueron descartadas para la construcción del mapa. La eliminación de estas bandas supone que el 39,8% de la variabilidad detectada por la metodología AFLP no superó alguno de los requisitos de intensidad, definición, reproducibilidad y/o ajuste a la segregación mendeliana aplicados a la selección de marcadores. Particularmente, el porcentaje de marcadores desviados significativamente ($P < 0,05$) de las proporciones esperadas de segregación fue significativamente mayor en los AFLP (25,7%) que en los microsatélites (4,1%). Esta diferencia entre marcadores aplicados a una misma muestra es más indicativa de errores de genotipado en los AFLP, que de factores biológicos subyacentes a la desviación. En comparación con la tecnología microsatélite, la mayor sensibilidad de la metodología AFLP a las condiciones de calidad y concentración del ADN de partida, así como la complejidad de su protocolo que incluye etapas de restricción, ligamiento y amplificación de secuencias anónimas, podría introducir cierta variabilidad derivada de artefactos que afectarían a la reproducibilidad de los perfiles AFLP. Asimismo, la fuerte distorsión de segregación detectada en algunos fragmentos AFLP en los que los genotipos de ausencia de banda fueron casi anecdóticos (7 de 46 en E32/T49.335; 5 de 45 en E32/T49.421; ver Tabla III.7) hace pensar que se trata de bandas monomórficas, que no fueron amplificadas o detectadas en ciertos individuos. Alguno de los factores técnicos expuestos a continuación podrían ser los responsables de estos falsos negativos.

Cantidad y calidad del ADN de partida. En algunos fragmentos polimórficos detectados en el presente estudio, se observaron diferencias

notables en la intensidad de banda entre los individuos de fenotipo positivo. Esta disparidad podría deberse a diferencias en la calidad y cantidad del ADN extraído a partir de embriones haploides poco desarrollados. Si bien, no existe un acuerdo en lo referente a la cantidad de ADN ideal para el desarrollo del protocolo AFLP y se ha descrito la obtención de marcadores AFLP a partir de 50 ng (Liu *et al.*, 2003) hasta 700 ng de ADN (Papa *et al.*, 2005), se reconoce que la calidad y abundancia del material genómico de partida podrían afectar a la reproducibilidad de estos marcadores (Vos *et al.*, 1995). Los 100 ng que se utilizaron en este estudio coinciden con la cantidad utilizada por Wilson *et al.* (2002) en el desarrollo de marcadores AFLP para la construcción de un mapa de ligamiento en camarón (*Penaeus monodon*). Sin embargo, en el presente trabajo, la cantidad de ADN de partida vino condicionada por la naturaleza haploide de las muestras. La observación al microscopio de los embriones de rodaballo enviados desde el I.E.O de Vigo, reveló individuos con un desarrollo detenido en la fase de neurulación. Esto suponía extraer el ADN a partir de un número escaso de células encerradas en una estructura embrionaria rica en material graso, que no había llegado a ser consumido. El protocolo de extracción optimizado que se puso a punto para este fin confirmó la escasez del ADN contenido en los embriones. Además, en un número importante de individuos, el ADN extraído tuvo que ser descartado al resultar sucio o fragmentado, posiblemente debido al tiempo transcurrido desde la muerte del embrión hasta el aislamiento de su ADN. A pesar de las precauciones tomadas, algunos de los embriones seleccionados podrían haber detenido su desarrollo en tiempos diferentes, rindiendo distintas cantidades de ADN en el proceso de extracción, lo que se traduciría en diferencias en la intensidad de banda cuando se comparasen los perfiles de individuos más o menos desarrollados.

Defectos de migración y pérdida de marcaje fluorescente. Independientemente de la calidad genómica, otro factor que podría afectar a la correcta interpretación de los genotipos sería el marcaje de los cebadores, ya

que la utilización de cebadores marcados de baja calidad resultaría en picos de baja intensidad o en completa ausencia de señal (Liu *et al.*, 1999; Yu y Guo, 2002; Li y Guo, 2004).

Además de los fragmentos con segregación fuertemente distorsionada ($P < 0,01$), existen otros en los que la desviación de la segregación mendeliana fue más moderada ($0,01 < P < 0,05$). Sin embargo, en ambos casos, la desviación se debió mayoritariamente a un exceso de presencia de banda frente a ausencia. Algunos autores (O'Hanlon y Peakall, 2000; Vekemans *et al.* 2002; Bonin *et al.* 2004) postulan dos fenómenos como principales responsables de la desviación hacia un exceso de fenotipos positivos: amplificaciones inespecíficas y homoplasia de tamaño.

Amplificaciones inespecíficas. Cuando se compara el porcentaje de bandas que no superaron los requisitos de calidad entre los fragmentos amplificados por combinaciones de cebadores parecidos y los amplificados por combinaciones bien diferenciadas, el porcentaje de las primeras supera en más del doble al de las segundas (66,6% y 28,45%, respectivamente). Igualmente, el porcentaje de polimorfismos coincidentes en tamaño fue mayor entre las bandas amplificadas por combinaciones parecidas (12,5%) que entre las bandas amplificadas por combinaciones bien diferenciadas (9,6%). Todo ello podría ser indicativo de la presencia de amplificaciones inespecíficas resultantes de cierta permisividad en el anillamiento de cebadores que no se diferencian en más de una base. Estas amplificaciones introducirían ruido de fondo en los perfiles AFLP, dificultando especialmente la interpretación de los genotipos negativos, donde las bandas inespecíficas poco intensas y definidas se harían más evidentes pudiendo valorarse como falsos positivos. Pocos trabajos hacen referencia a esta posible fuente de error, si bien en la misma presentación de la metodología AFLP, sus autores reconocían que cierto nivel de amplificación inespecífica ocurriría siempre (Vos *et al.*, 1995) y el hecho de que estas bandas alcanzaran el nivel de detección dependería de las

condiciones de reacción. En un trabajo más reciente Rajapakse y colaboradores (2001) advertían de la problemática asociada al uso de cebadores con secuencias poco diferenciadas (por ejemplo, la elección de bases selectivas AAG y AAC), ya que podrían amplificar los mismos fragmentos genómicos. Para evitar tales duplicaciones, estos autores aconsejan seleccionar combinaciones de cebadores bien diferenciadas (por ejemplo, AAG vs CGT). Esta selección no siempre se pudo llevar a cabo en el presente trabajo, puesto que se partía de un conjunto limitado de cebadores.

Homoplasia de tamaño. Otro factor técnico que podría alterar las proporciones reales de segregación sería la homoplasia (Peakall *et al.*, 1998). Este fenómeno está relacionado con el sistema de detección de fragmentos AFLP, que asume sin verificación a nivel de secuencia, que los fragmentos de igual tamaño son homólogos, es decir, proceden de un mismo locus. Los mismos autores de la metodología, al comentar que “la mayoría de los fragmentos AFLP corresponden a posiciones únicas en el genoma” (Vos *et al.*, 1995) asumen, indirectamente, la existencia de una minoría no homóloga entre los fragmentos de igual tamaño. Este problema ya se planteó en otra de las técnicas que producen perfiles multilocus, los RAPDs, cuando autores como Clark y Lanigan (1993) sugirieron que debía testarse sistemáticamente la homología de las bandas de igual tamaño. En el presente estudio, la identificación de fragmentos complejos dentro del perfil AFLP amplificado por los cebadores E33/T49, puso en evidencia la relevancia de este fenómeno en la alteración de las proporciones de segregación de ciertos marcadores, cuyos genotipos resultan de la suma de loci independientes. Entre las 10 bandas polimórficas identificadas por la combinación E33/T49, la homoplasia parece afectar a sólo una de ellas, concretamente a una de las tres bandas que presentaron distorsión de segregación y que resultaría de la suma de dos loci diferentes: uno cuyo nucleótido adyacente al extremo 3' de *TaqI*-CAG sería una A; y otro en el que el nucleótido adyacente a dicho extremo sería una C. Si bien debería estudiarse el fenómeno en otras combinaciones de cebadores, a

partir de los resultados disponibles parece que la incidencia de la homoplasia es relativamente baja (afectaría al 10% de los polimorfismos E33/T49). Además, la ulterior comprobación del ajuste de la segregación a las proporciones mendelianas descartaría otras bandas complejas que pudieran formarse.

Considerando la bibliografía, estudios de mapeo comparativo han confirmado el escaso efecto de la homoplasia al encontrar que los marcadores AFLP son, predominantemente, reproducibles y locus-específicos (Haanstra *et al.*, 1999; Vuylsteke *et al.*, 1999; Pèrin *et al.*, 2002). Más concretamente, Vekemans y colaboradores (2002) cuantificaron el efecto de la homoplasia a través de simulaciones teóricas, obteniendo que, bajo dicho fenómeno, la proporción de fragmentos detectables en un perfil AFLP sería del 88%, mientras que el 12% restante correspondería a bandas coincidentes en tamaño que se asumirían como loci únicos. Los mismos autores encuentran que la asunción teórica del fenómeno homoplásico reproduce la típica distribución no uniforme de los perfiles AFLP, con una tendencia a decrecer la frecuencia de fragmentos a medida que aumenta su tamaño y una mayor presencia de bandas monomórficas entre los fragmentos más pequeños. El incremento de la proporción de bandas detectables con el tamaño de las mismas indica un menor efecto de la homoplasia entre los fragmentos de mayor tamaño, por lo que los autores recomiendan evitar el uso de las bandas cortas, a fin de disminuir el riesgo de homoplasia. En el presente estudio, el porcentaje de marcadores AFLP que no superaron las 300 pb es del 62,8%, siendo el 50% de estos marcadores menores de 200pb. Si bien la mayoritaria presencia de fragmentos de moderado tamaño podría alertar sobre una mayor incidencia de homoplasia, el porcentaje de fragmentos complejos testados en el perfil AFLP E33/T49 no superó el 12% descrito por Vekemans *et al.* (2002).

IV.2.3. Algunas consideraciones acerca de los marcadores microsatélite y AFLP en el análisis de ligamiento.

La metodología AFLP, gracias a la exploración simultánea de múltiples regiones del genoma, resulta más rápida en la identificación de marcadores que la microsatélite. En este trabajo, se pudieron identificar hasta 10 marcadores a partir de una sola combinación de cebadores. Considerando tal productividad, los AFLP parecen ser los marcadores más adecuados para la construcción rápida de un mapa genético denso. Sin embargo, tal como señalaban Kocher *et al.* (1998) al discutir la idoneidad de los AFLPs en la construcción del mapa genético de tilapia “la velocidad tiene un precio”, que con frecuencia se traduce en una mayor incidencia de errores de genotipado. Estos errores unidos a las propias características de la metodología AFLP dificultan su transferibilidad entre laboratorios, cruzamientos y especies. Si bien varios estudios en plantas han utilizado con éxito los marcadores AFLP para relacionar mapas independientes (Waugh *et al.*, 1997; Alonso-Blanco *et al.*, 1998; Qi *et al.*, 1998; Rouppe van der Voort *et al.*, 1998; Haanstra *et al.*, 1999; Vuylsteke *et al.*, 1999), la distancia de migración entre fragmentos no parece una información suficiente para identificar de manera inequívoca un determinado locus en diferentes poblaciones de mapeo (Plomion *et al.*, 1995). En este sentido, los marcadores codominantes, multialélicos y específicos de secuencia, tales como los microsatélite, serían los más consistentes para la comparación entre mapas (Cervera *et al.*, 2001; Rajapakse *et al.* 2001; Neale *et al.*, 2002; Pèrez *et al.*, 2004). En proyectos de mapeo modestos, como el presentado en esta memoria, unos pocos marcadores microsatélite bien distribuidos entre los grupos de ligamiento (LGs), constituyen valiosos puntos de anclaje para establecer la correspondencia entre los grupos de ligamiento de mapas desarrollados con diferentes familias de la misma especie o incluso de especies próximas (Neale *et al.*, 2002).

Considerando la transferibilidad de los loci microsatélite entre distintos linajes y especies, a medida que progresan los proyectos de mapeo genómico, aumenta el interés en estos marcadores en cuanto a su potencial para proporcionar datos comparativos de posibles homologías/ortologías entre loci de interés (Walter *et al.*, 2004). En esta misma línea, las propias genotecas desarrolladas para la caracterización de nuevos loci microsatélite, se han revelado como una interesante fuente de información para el estudio de las relaciones evolutivas interespecíficas. Esta información se basa en la comparación de las secuencias flanqueantes que se insertan en los plásmidos acompañando a los loci microsatélite, con las secuencias nucleotídicas derivadas de los proyectos de secuenciación a gran escala en especies modelo (Streelman *et al.*, 1998). Estas secuencias tienen el potencial de proporcionar información en cuanto a la sintenia entre distintos taxones, de modo que las “costosas” genotecas de microsatélites resultan de gran interés informativo para la genómica comparativa. Así, cuando se abarcan proyectos de mapeo, incluso de moderado alcance como el presentado en esta memoria, es interesante la inversión en este tipo de herramientas génicas, especialmente si se consideran futuras aplicaciones de estos mapas.

En cuanto a la función de aplicación más inmediata de los mapas de ligamiento, es decir, la identificación de loci implicados en caracteres cuantitativos (QTL), los marcadores multialélicos, como los microsatélite, resultan más útiles que los bialélicos dominantes AFLPs, puesto que permiten estimar con mayor precisión los efectos aditivos y dominantes de los QTLs (Sewell *et al.*, 2000; Chagnè *et al.* 2002). Sin embargo, a pesar de sus limitaciones, los AFLPs se han usado con éxito en la construcción de mapas de ligamiento en plantas para el análisis de QTLs relacionados con caracteres agronómicos, tales como resistencia a enfermedades y tolerancia a la sal (Quarrie *et al.*, 1997; Voorrips *et al.*, 1997; Jin *et al.*, 1998; Rouppe Van Der Voort *et al.*, 1998) y para identificar QTLs asociados con caracteres fisiológicos en roedores (Otsen *et al.*, 1996). A pesar de sus limitaciones, la capacidad de

la tecnología AFLP para ensayar simultáneamente cientos de loci en cualquier genoma es una de las características que siguen haciendo atractiva a esta técnica para la búsqueda de los genes implicados en caracteres complejos (Ajmone-Marsan *et al.*, 1997).

IV.3. MAPA GENÉTICO DE *S. maximus*

IV.3.1. Ligamiento genético entre pares de marcadores.

Los 23 grupos de ligamiento detectados en el presente mapa de rodaballo parecen una buena aproximación al número haploide de cromosomas de la especie, $n = 22$. Sin embargo, el mapa genético se basa en una representación probabilística cuya consistencia depende fundamentalmente del umbral LOD utilizado para el análisis de ligamiento entre pares de marcadores.

En este mapa se utilizó un LOD mínimo de 3,0 para aceptar ligamiento entre los marcadores del mapa *framework*. Esta puntuación LOD es el umbral recomendado en la mayoría de los estudios para la construcción de mapas (Maliepaard *et al.*, 1997). La utilización de valores LOD más bajos aumentaría la probabilidad de cometer errores de tipo I, si bien no hay reglas generales, y la inclusión de falsos positivos dependerá del tipo de datos utilizados (Ellis *et al.*, 1992). En estudios con un número relativamente bajo de marcadores, un LOD 3 puede resultar demasiado severo para detectar ligamiento y en estos casos se suele recurrir a un LOD 2,0 más permisivo (Paglia *et al.*, 1998; Cervera *et al.*, 2001; Pérez *et al.*, 2004). En el presente estudio, la utilización de un LOD mínimo 2,0 permitió identificar tres dobletes entre los marcadores no ligados a LOD 3,0, así como la asignación de nuevos marcadores a los grupos de referencia *framework* preestablecidos. En todos los casos, estos marcadores se situaron en los extremos de los grupos de ligamiento, lo que sería indicativo de una elevada frecuencia de recombinación respecto a los marcadores *framework*, ya sea por una distancia real entre marcadores o por la

presencia de posibles errores de genotipado (Collins *et al.*, 1996; Li *et al.*, 2005).

Los marcadores que resultaron no ligados a LOD 2,0 corresponderían a loci alejados en los extremos de los cromosomas, que no llegan a alcanzar la puntuación umbral para la detección significativa de ligamiento (Lander *et al.*, 1987). Cuando faltan marcadores intermedios, la capacidad de detección de asociación entre marcadores podría resultar afectada por el tamaño muestral (Plomion *et al.*, 1995; Coimbra *et al.*, 2003; Moen *et al.*, 2004). En el presente estudio, el número de descendientes analizados (50) está dentro del rango utilizado en trabajos similares (Kocher *et al.*, 1998; Gates *et al.*, 1999; Sung y Liang, 2004), sin embargo, como ya se expuso en Material y Métodos, la potencia estadística de dicho tamaño muestral para estimar frecuencias de recombinación entre marcadores alejados (frecuencia de recombinación ($r > 30\%$) o muy próximos ($r < 7\%$) es limitada (Lie *et al.*, 1994). Según estas previsiones, los marcadores no ligados presentarían frecuencias de recombinación por encima del 30% respecto a los marcadores mapeados.

La falta de cobertura en ciertas regiones genómicas se manifiesta también por la presencia de un grupo de ligamiento extra del mapa framework (23 LGs) en relación al número haploide de cromosomas de la especie ($n=22$, Bouza *et al.*, 1994a). Este exceso de grupos es habitual en los mapas preliminares con una densidad relativamente baja de marcadores (Postlethwait *et al.*, 1994; Wada *et al.*, 1995; Kocher *et al.*, 1998), donde los loci más distantes en el cromosoma son identificados como LG pequeños e independientes (revisión de Danzmann y Gharbi 2001; Pérez *et al.*, 2004).

Predecir el tamaño muestral para combinar los LGs y los marcadores no ligados en los 22 grupos correspondientes al número haploide de cromosomas de la especie no es tarea fácil. Plomion *et al.* (1995) encontraron que doblando el número de individuos analizados apenas lograban un ligero incremento en la precisión de las posiciones relativas de los marcadores

accesorios, mientras Kesseli *et al.* (1994) no consiguieron reducir el número de LG extra, a pesar de doblar el número de marcadores utilizados. Las características propias del genoma a analizar, tales como frecuencia de recombinación y tamaño genómico, así como la calidad de los propios marcadores influirán en la mayor o menor eficacia a la hora de detectar ligamiento.

Dadas las características del genoma de rodaballo, pequeño y con un número relativamente bajo de cromosomas, quizá la estrategia más adecuada para aumentar la confianza del análisis de ligamiento dos-puntos sería incrementar, no tanto el número de meiosis analizadas, sino el número y la calidad de los marcadores, lo que permitiría corroborar la consistencia del mapa aplicando umbrales de detección LOD más altos (Chakravarti *et al.*, 1991). Aún así, la discrepancia LG vs n puede mantenerse, incluso en mapas consolidados con elevado número de marcadores codominantes (Lee *et al.*, 2005).

IV.3.2. Ordenación de los marcadores en los grupos de ligamiento.

En el presente estudio, los mapas de cada uno de los LG con tres o más marcadores se apoyan en una puntuación LOD al menos dos unidades mayor a la del orden alternativo más probable, lo que significa que la probabilidad de que las observaciones procedan del mapa presentado es 100 veces mayor que la probabilidad de que surjan del mapa alternativo (Lander y Bostein, 1989).

En los grupos con mayor número de marcadores, el mapa de cada LG lo constituye un núcleo de 4 o 5 loci ligados con una alta puntuación LOD ($\geq 3,0$), mientras, que los loci con una menor puntuación para el ligamiento se sitúan en los extremos. Estos loci periféricos, en ocasiones, aparecen separados de los loci más próximos por largos intervalos (*gaps*),

particularmente elevados (> 25 cM) en los LGs 5 y 11. En mapas densos como el presentado por Knapik *et al.* (1998) para el pez cebra con una resolución media de 3,3 cM se mantenían *gaps* de entre 20 y 30 cM, que los autores explicaban como resultado del azar en la distribución de los marcadores mapeados. En otros casos, la presencia de largos *gaps* en los mapas se ha justificado como posibles puntos calientes de recombinación en el genoma (Paglia *et al.*, 1998; Saal y Wricke, 2002) o bien como regiones no cubiertas por los marcadores utilizados, ya sea por la inaccesibilidad a las enzimas de restricción empleadas o por la menor frecuencia de secuencias repetitivas (Cervera *et al.*, 2001; Saal y Wricke, 2002). Dada la densidad del presente mapa, el simple azar podría justificar la presencia de los *gaps* observados, puesto que es posible que regiones amplias del genoma no hayan sido cubiertas por los marcadores (Vallejos *et al.*, 1992). El ajuste de los marcadores observados a la distribución de Poisson apoya esta posibilidad. Sería necesario aumentar la densidad del mapa para aproximarse a la naturaleza de estos intervalos genómicos. En futuros estudios y a medida que se disponga de mapas más densos de rodaballo, sería interesante abordar la asignación de los grupos de ligamiento a los cromosomas y comparar la distribución de los *gaps* con los patrones de bandeo C descritos en estudios citogenéticos de esta especie (Bouza *et al.*, 1994a; Pardo *et al.*, 2001).

Algunos LGs, además de los marcadores *framework*, incluyen también marcadores accesorios, cuya posición no pudo determinarse bajo el criterio de ordenación establecido, a pesar de que el análisis dos-puntos hubiera detectado firmes evidencias de su ligamiento a dichos LGs. En el caso de los marcadores E33/T37.352r, 2/5TG14r y 5/4CA20 accesorios en el LG3, así como de E32/T50.236 en el LG15, el LOD de asociación al marcador *framework* más próximo fue superior a 3,0 y la distancia genética inferior a 7 cM. Según los cálculos de Lie *et al.* (1994), el estrecho ligamiento entre marcadores dificultaría la localización precisa de los mismos, dado el limitado número de meiosis analizadas en la muestra (Grattapaglia y Sederoff, 1994;

Plomion *et al.*, 1995; Kocher *et al.*, 1998). Sin embargo, esta explicación no es aplicable a los demás marcadores accesorios de este estudio, que mayoritariamente distan más de 20 cM de su marcador *framework* más próximo y se ligan con un LOD no superior a 2,0. La alta frecuencia de recombinación y la baja puntuación para la asociación podrían ser indicativas de errores de genotipado o de falta de datos en el conjunto muestral, lo que dificultaría la ordenación de estos loci (Grattapaglia y Sederoff, 1994; Wilson *et al.*, 2002). Dado que el porcentaje de meiosis informativas superó el 70% en todos los marcadores genotipados, es probable que los errores experimentales más que la falta de datos hayan interferido en la ambigua localización de estos marcadores. Posiblemente, la inclusión de nuevos loci y el consecuente incremento en la densidad del mapa ayudará a esclarecer las relaciones de ligamiento de estos marcadores, precisando su localización en las regiones distales.

Otro aspecto interesante que pone de manifiesto la ordenación de marcadores en el mapa es la presencia de marcadores AFLP codominantes. Entre los AFLPs mapeados se identificaron cinco pares de marcadores completamente ligados en fases de ligamiento opuestas y cuyos tamaños moleculares difieren en pocos pares de bases. Estos marcadores suponen el 4,6% del total de AFLP seleccionados para el análisis de ligamiento, porcentaje superior al 2,5% descrito en la platija japonesa (Coimbra *et al.*, 2003), pero muy próximo al 5% obtenido en tilapia (Kocher *et al.*, 1998). Estas secuencias podrían proceder de variaciones en la longitud de los fragmentos debido a pequeñas inserciones o deleciones, generalmente en regiones repetitivas, dando lugar a alelos de diferente tamaño. El aislamiento y secuenciación de los fragmentos de ADN correspondientes a este tipo de marcadores, permitiría confirmar el alelismo entre las bandas AFLP y al mismo tiempo, podría llevar a la identificación y caracterización de nuevos marcadores microsatélite.

En la misma línea, la ordenación de los marcadores en los LGs reveló la cosegregación de ciertos pares de marcadores microsatélite tan próximos en el genoma que a efectos de mapeo funcionan como un único marcador (F8-ICA26 y F8-I11/8/17r, F12I-AG18r y F12I-TG, Iyengar *et al.*, 2000; Smax-04 y Smax-04b, Coughlan *et al.*, 1996; Bouza *et al.*, 2002). El ligamiento completo es congruente con la estrecha proximidad física de estos marcadores, que según describen sus autores fueron aislados a partir de los mismos clones procedentes de librerías enriquecidas (Iyengar *et al.*, 2000; Bouza *et al.*, 2002).

IV.3.3. Distancia de mapa.

Implícito en la ordenación de los marcadores está el cálculo de distancias genéticas. El programa MAPMAKER ordena los marcadores en función de las frecuencias de recombinación multipunto y entonces aplica una función de mapa (Haldane o Kosambi) para el cálculo de distancias (Lander *et al.*, 1987).

Entre la variedad de funciones de mapa que se han desarrollado (Haldane, 1919; Kosambi, 1944; Carter y Falconer, 1951; Karlin, 1984), no parece haber una única y mejor función de mapa, sino que para un conjunto de datos deberá elegirse la función que mejor se adapte, considerando, fundamentalmente, la interferencia cromosómica de la especie en estudio o de especies filogenéticamente próximas. Unas y otras funciones asumen diferentes niveles de interferencia para corregir la fracción observada de recombinación considerando los múltiples entrecruzamientos no observados (Windemuth *et al.*, 1998).

En rodaballo no hay datos previos acerca del grado de interferencia de su genoma, si bien, datos no publicados parecen indicar una interferencia elevada (P. Martínez, comunicación personal). El nivel de interferencia elevado coincidiría con la tendencia observada en otras especies de Teleósteos, que con frecuencia se ha caracterizado por altos niveles de interferencia en

comparación con otros animales o plantas (Thorgaard *et al.*, 1983). La interferencia elevada es común al pez cebra (Johnson *et al.*, 1996), medaka (Naruse y Shima, 1989), pez gato (Liu *et al.*, 1992) y tilapia (Kocher *et al.*, 1998). Algunos autores relacionan el alto grado de interferencia con el tamaño relativamente pequeño de los cromosomas de peces, que podría inhibir la recombinación y permitir un sólo entrecruzamiento por brazo cromosómico (Kocher *et al.*, 1998). En este contexto, la función de mapa Kosambi, que asume un incremento de la interferencia con el aumento de la distancia entre marcadores (Kosambi 1944), es una de las funciones más utilizadas para calcular las distancias genéticas en los mapas de peces.

IV.3.4. Distribución de los marcadores.

Se espera que los marcadores AFLP y microsatélite se distribuyan al azar cuando los polimorfismos están distribuidos al azar (Cervera *et al.*, 2001). Esta suposición se evaluó dentro de cada grupo de ligamiento y a lo largo de todo el mapa de rodaballo.

Distribución de marcadores dentro de los grupos de ligamiento

A nivel de grupos de ligamiento particulares, la distribución aleatoria de los marcadores *framework* es indicativa de que los polimorfismos subyacentes a estos marcadores se reparten al azar en los cromosomas de rodaballo. Sin embargo, la inclusión de los marcadores accesorios en el análisis reveló en los grupos más densos (LG1 y LG3) la existencia de pequeñas regiones donde la concentración de marcadores superó la esperada por azar. En los mapas más saturados la formación de estas agrupaciones es un suceso muy común, independientemente del organismo o del tipo de marcador utilizado (Grattapaglia y Seredoff, 1994). En la literatura, se han esgrimido dos argumentos para justificar la presencia de dichas agrupaciones: (1) Las distancias genéticas, basadas en frecuencias de recombinación, podrían no reflejar la localización física de los marcadores. En este sentido, las concentraciones de marcadores podrían corresponder a las regiones

centroméricas donde la recombinación está prácticamente suprimida (Tanksley *et al.*, 1992; Young *et al.*, 1998); (2) Alternativamente, los marcadores podrían estar físicamente concentrados a lo largo del cromosoma, ya sea por una distribución desigual de los polimorfismos o por la presencia de secuencias repetitivas, especialmente elevada en las áreas pericentroméricas (Alonso-Blanco *et al.*, 1998; Qi *et al.*, 1998; Lidner *et al.*, 2000).

El número de marcadores microsatélite mapeados en el presente estudio, fue mucho menor que el de AFLPs, lo cual pudo limitar la potencia estadística para la detección de desviaciones en su distribución. De hecho, los agrupamientos de marcadores identificados en este mapa de rodaballo, incluyen exclusivamente marcadores de tipo AFLP. La difusión de la tecnología AFLP reveló una especial tendencia de estos marcadores a la agrupación en las regiones centrales de los LG, apreciable no sólo en los mapas más densos, sino también en mapas de densidad moderada. Particularmente en peces, se han descrito agrupaciones de marcadores AFLP en tilapia, en la trucha arcoiris y en medaka, si bien el nivel de agrupamiento parece aumentar con la densidad del mapa. Así, con 112 AFLPs se identificaron un par de agrupamientos en el mapa de tilapia (Kocher *et al.*, 1998); con 219 AFLP se observaron más de cinco agrupamientos en el mapa de la trucha arcoiris (Robinson *et al.*, 2001); y numerosas concentraciones de AFLPs al mapear 488 marcadores en medaka (Naruse *et al.*, 2000). En el presente estudio, la posición que ocupan las agrupaciones de AFLP, ligeramente descentradas en los LG más densos, apunta a la posibilidad de que estén identificando los centrómeros de algunos de los cromosomas subtelo/acrocéntricos de rodaballo. Sin embargo, la baja densidad manejada en este mapa, así como la existencia de otras regiones cromosómicas de heterocromatina intersticial no centromérica, impone prudencia en la lectura de estas interpretaciones.

Dado que la concentración de marcadores podría reflejar un sesgo en la composición de bases de las regiones mapeadas, la propia composición

nucleotídica de los cebadores AFLP podría influir en la distribución de este tipo de marcadores (Paglia *et al.*, 1998). Así, algunos autores han sugerido el uso del par enzimático *EcoRI/PstI* frente al tradicional *EcoRI/MseI*, cuya secuencia rica en AT ha demostrado una tendencia a la amplificación de regiones repetitivas centroméricas o pericentroméricas y por tanto, a la formación de agrupaciones de marcadores en los mapas (Roupe Van Der Voort *et al.*, 1998; Castiglioni *et al.*, 1999; Vuylsteke *et al.*, 1999). Los marcadores AFLP derivados de la digestión con *EcoRI/PstI* han demostrado una distribución más equilibrada a lo largo del genoma. Sin embargo, este par de enzimas presenta sensibilidad a la metilación, lo que, si bien es positivo a la hora de identificar posibles secuencias codificadoras, limita considerablemente la cantidad de marcadores identificados (Castiglioni *et al.*, 1999; Vuylsteke *et al.*, 1999; Saal y Wricke, 2002). En el caso de la combinación de endonucleasas *EcoRI/TaqI*, la composición equilibrada de bases en la diana del enzima *TaqI*, evitaría el sesgo hacia la amplificación de secuencias repetitivas. Consecuentemente, una concentración de marcadores derivados de este par enzimático apoyaría la reducción de la recombinación como causa de los agrupamientos y por tanto, facilitaría la identificación de regiones centroméricas en los mapas. En todo caso, se necesitarían mapas AFLP más densos basados en esta combinación *EcoRI/TaqI* en otras especies para establecer con mayor consistencia el patrón de distribución de estos marcadores.

Con la aplicación de los test estadísticos de distribución de marcadores se contempló también el fenómeno opuesto a la concentración de marcadores, es decir, la presencia de largos intervalos o *gaps* entre los algunos loci. Sin embargo, en ningún caso la presencia de estos intervalos fue mayor de la esperada por azar.

Distribución de marcadores entre los grupos de ligamiento

La presencia de agrupaciones de marcadores en determinados intervalos del mapa, también se puso de manifiesto en el estudio de la

distribución de marcadores entre los grupos de ligamiento, al determinar que la frecuencia observada de intervalos con 5 marcadores superaba la esperada por azar. Al mismo tiempo, estas concentraciones de marcadores conllevan un aumento del número de intervalos con un solo marcador, cuya presencia también resultó ligeramente superior a la esperada según la distribución de Poisson. Sin embargo, cuando se estudió la relación entre el número total de marcadores y el tamaño en cM del grupo de ligamiento en que mapean, se encontró una correlación positiva, indicativa de la aleatoriedad en la distribución (Cervera *et al.*, 2001; Pérez *et al.*, 2004).

Cuando en el estudio de la distribución se consideró no sólo el número de marcadores, sino su origen o composición, se encontró que en algunos LGs, la concentración de marcadores AFLP procedentes de la misma combinación de cebadores superaba la esperada por azar ($P < 0,05$). Este fenómeno se ha observado en otros mapas genéticos de peces basados mayoritariamente en marcadores AFLP (Kocher *et al.*, 1998; Pérez *et al.*, 2004). Algunos autores proponen el alelismo entre las bandas AFLP como posible causa de la agrupación entre estos marcadores (Alonso-Blanco *et al.*, 1998). Este sería claramente el caso de los marcadores E33/T37.217 y E33/T37.206r, que fueron identificados como alelos codominantes en el LG3 tras el análisis de ligamiento. Sin embargo, el mismo fenómeno queda descartado a la hora de justificar la agrupación en la misma región de otros dos marcadores de dicha combinación de cebadores, los E33/T37.352r y E33/T37.356r, dado que ambos marcadores segregan en la misma fase de ligamiento. Como ya se argumentó anteriormente, la causa subyacente a esta agrupación podría estar en la presencia de secuencias repetitivas, si bien, sería necesario aislar y secuenciar los fragmentos ALFP para confirmar esta hipótesis. Otro de los agrupamientos se encuentra en el LG2, donde tres marcadores de la combinación E42/T37 aparecen ligados, dos como loci *framework* y un tercero ligado a ellos como marcador accesorio. En este caso, las diferencias de tamaño entre los marcadores, el hecho de que segreguen en

la misma fase de ligamiento y las elevadas frecuencias de recombinación entre ellos descartan un posible alelismo entre algún par de estos marcadores. Un sesgo en la presencia de secuencias complementarias a la combinación de cebadores E42/T37 en el cromosoma correspondiente al LG2 parece la causa más probables de la asociación.

En ausencia de aparente ligamiento a alelos letales o semiletals, cuyo efecto sería presumiblemente bajo en embriones de edad temprana, algunos autores justifican la presencia de agrupaciones de marcadores AFLP originados por una misma combinación de cebadores, como el resultado de errores de genotipado (Kocher *et al.*, 1998; Pérez *et al.*, 2004). Particularmente, cuando en el agrupamiento están implicados marcadores con ratios de segregación distorsionados. Problemas en la amplificación que lleven a una disminución del número de genotipos disponibles o la presencia de amplificaciones inespecíficas o de homoplasia de tamaño, que homogenicen algunos de los patrones de bandas de las combinaciones de cebadores afectadas, podrían llevar a la interpretación de falsos ligamientos (Faris *et al.*, 1998; Virk *et al.*, 1998; Negi *et al.*, 2000). Éste podría ser el caso de los marcadores AFLP procedentes de las combinaciones E33/T38 y E33/T58 cuya presencia es mayor de la esperada en los LGs 5 y 8, respectivamente. En ambos casos, en las agrupaciones están implicados marcadores con razones de segregación distorsionadas. Además, son combinaciones de cebadores en las que, por dificultades en la amplificación, se disponía de pocas meiosis, lo que podría haber contribuido a la distorsión y disminuiría la confianza de su localización en el mapa.

En cuanto a los marcadores microsatélite, el escaso número de marcadores mapeados en este estudio limita la detección de desviaciones de la distribución aleatoria, por lo que es preciso tener cautela en la interpretación de los resultados. En todo caso, entre los 20 loci microsatélite mapeados ninguno se vio implicado en alguna de las concentraciones de marcadores, ni

su número superó el esperado en ninguno de los 10 LGs en los que se distribuyeron. Esta distribución aleatoria y equilibrada de los microsátélites coincide con lo descrito en la literatura (Dietrich *et al.*, 1996), donde con frecuencia aparecen con una distribución distal en los mapas menos densos, lo que evita su implicación en las supuestas concentraciones pericentroméricas (Young *et al.*, 1998).

IV.3.5. Distorsión de segregación.

Dado un determinado umbral LOD, algunos autores han descrito diferencias en el porcentaje de marcadores mapeados según su mayor o menor ajuste a las proporciones de segregación mendelianas (Li y Guo, 2004), siendo menor el porcentaje de marcadores asignados al mapa entre aquellos con razones de segregación distorsionadas.

En el presente mapa no se apreciaron tales diferencias. Contrariamente, el porcentaje de marcadores con segregación distorsionada asignados al mapa (87,5%) fue muy similar al de marcadores no distorsionados (86,7%). Sin embargo, se dieron diferencias en la puntuación LOD con la que unos y otros marcadores se incluyeron en el mapa. Mientras que el 75% de los marcadores con segregación mendeliana se unen al mapa *framework* ($\text{LOD} \geq 3,0$), tan sólo el 60% de los marcadores con ratios distorsionados se ligan en estas circunstancias. Esta diferencia da una idea de la dificultad que puede presentar el mapeo de marcadores distorsionados, puesto que aumenta el riesgo de cometer errores de tipo I de falso ligamiento (Gerber y Rodolphe, 1994; Cloutier *et al.*, 1997; Cervera *et al.*, 2001). Sin embargo, se recomienda no excluirlos por completo de la construcción de los mapas genéticos. La desviación en estos marcadores de la razón de segregación esperada podría ser el resultado de una asociación con genes deletéreos (Verhaegen y Plomion, 1996; Kuang *et al.*, 1999). Si se rechazasen los marcadores desviados, las regiones genómicas que incluyesen a estos genes quedarían sin mapear, perdiendo además marcadores que podrían ser valiosos para

programas de selección asistida por marcadores (MAS: *Marker Assited Selection*). Considerando esta posibilidad, en el presentes estudio, se incluyeron en el análisis de ligamiento los marcadores que se desviaron al nivel del 5%, pero no al 1%, y sólo se mapearon una vez establecida la posición de los marcadores *framework*.

A propósito de la asociación de los marcadores a genes deletéreos, varios trabajos describen una tendencia de los marcadores distorsionados a agruparse en ciertos grupos de ligamiento (Young *et al.*, 1998; Hurme y Savolainem, 1999; Li y Guo, 2004). En algunos casos se han podido identificar asociaciones entre estos marcadores y genes de resistencia a enfermedades (Cervera *et al.*, 2001; Launey y Hedgecock, 2001). En el presente mapa de rodaballo, no se han identificado asociaciones consistentes que sugieran bases biológicas para la desviación. Sin embargo, en tres de los grupos de ligamiento mayores (LG1, LG4 y LG5), con al menos dos marcadores distorsionados, éstos se dispusieron en posiciones adyacentes. Dada la corta edad de los embriones y la diferencia en el porcentaje de marcadores distorsionados entre AFLPs (25%) y microsatélites (4,1%), posibles errores de genotipado en la caracterización de los AFLPs, junto con el relativamente pequeño tamaño muestral, parecen la opción más conservativa para explicar la distorsión.

En todo caso, el porcentaje de marcadores AFLP distorsionados es muy inferior al descrito en otras publicaciones para este tipo de marcador (65% en el mapa de la col (Voorrips *et al.*, 1997); y 54% en el del gusano de seda (Tan *et al.*, 2001)), lo que sugiere que la incidencia de errores en el genotipado fue moderada.

IV.3.6. Longitud genómica y cobertura del mapa

El tamaño genómico, en unidades de mapa o centimorgans (cM), es una de las características fundamentales de los mapas genéticos de cada especie (Chakravarti *et al.*, 1991). En especies en las que existen mapas

saturados que cubren casi la totalidad del genoma, la longitud de mapa puede obtenerse como la suma de todos los intervalos mapeados, tal es el caso de *Arabidopsis* (Reiter *et al.*, 1992) o del pez cebra (Knapik *et al.*, 1998). Sin embargo, en mapas de baja o moderada densidad, el resultado de esta suma sería una infraestima del tamaño genético real y se suelen utilizar distintas aproximaciones que tienen en cuenta el número de grupos de ligamiento extra o de marcadores que segregan independientemente para obtener su estimación (Postelwaith *et al.*, 1994).

En el presente estudio, la suma de las distancias genéticas en los intervalos del mapa *framework* resultó en una longitud total de 790,6 cM, que aumenta hasta los 1010,4 cM, cuando se incluyeron en el cómputo los grupos formados por pares de marcadores ligados. Sin embargo, la presencia de marcadores no ligados y de un grupo de ligamiento extra en relación al número de cromosomas de la especie, indican que la longitud total del mapa es mayor de la cubierta (Moen *et al.*, 2004). A partir del mapa genético parcial y aplicando los estimadores de Chakravarti y Fishman (Chakravarti *et al.*, 1991; Fishman *et al.*, 2001) se obtuvo, en promedio, una longitud esperada de 1148,5 cM para el genoma de rodaballo. Esta estima, basada en frecuencias de recombinación de una hembra de rodaballo, es similar a la obtenida en otras especies de peces a partir de mapas parciales basados en AFLPs: tilapia (Kocher *et al.*, 1998); medaka (Naruse *et al.*, 2000); y particularmente, próxima a la longitud estimada para el mapa genético femenino en la platija japonesa, dentro del orden de los Pleuronectiformes (1176,4 cM, Coimbra *et al.*, 2003).

La precisión de estas estimas depende tanto del tamaño y tipo de población de mapeo, como del tipo y número de marcadores (Powell *et al.*, 1997; Pèrin *et al.*, 2002), si bien se recomienda aumentar el número de marcadores, más que el número de meiosis para mejorar la precisión de la estima (Chakravarti *et al.*, 1991). En el presente mapa, dado el porcentaje de marcadores no ligados, la probabilidad de que se unan nuevos marcadores es

del 84,7%, lo que sugiere que el mapa no se extenderá llamativamente con la inclusión de más marcadores, si bien la eventual adición de marcadores a las regiones teloméricas mejoraría la cobertura del mapa (Saal y Wricke, 2002; Huber y Hedgecock, 2004; Chistiakov *et al.*, 2005).

Tener una estima de la longitud genómica esperada (G_e) es fundamental para diseñar un experimento de ligamiento eficiente en cualquier especie, puesto que la determinación del número de marcadores necesarios para cubrir un genoma es críticamente dependiente del valor de G_e (Bishop *et al.*, 1983; Chakravarti *et al.*, 1991). En el actual mapa genético femenino de rodaballo, los marcadores *framework* cubren un 68,8% de la longitud genómica estimada. Este porcentaje coincide con la cobertura teórica estimada según el método de Lange y Boenke (1982), lo que es indicativo del reparto aleatorio y equilibrado de los marcadores a lo largo del mapa (Kesseli *et al.*, 1994; Echt y Nelson, 1997; Cervera *et al.*, 2001). El mismo método estima que serían necesarios unos 118 marcadores para cubrir un 85% de la longitud del genoma de rodaballo siendo la distancia máxima entre marcadores adyacentes de 20 cM. Este número coincide con la cobertura obtenida cuando se considera el mapa total (87,9%).

IV.3.7. Distancia física vs. distancia genética

Si bien 1 cM se define como un 1% de recombinación génica entre un par de marcadores, el número real de pares de bases incluidas en esta distancia depende del valor de la tasa de recombinación (Danzmann y Gharbi, 2001). Se ha demostrado que dicha tasa varía no sólo entre especies, sino entre familias e individuos dentro de una misma especie y que resulta afectada por el sexo, la secuencia de ADN y la estructura cromosómica (Robinson, 1996; Nachman y Churchill, 1996; Ohara *et al.*, 2005). Todo ello complica la traducción de las distancias genéticas en distancias físicas y la aplicación de los mapas para el marcaje de genes de interés y su eventual clonación posicional.

Diferencias interespecíficas en las tasas de recombinación. Entre las especies marinas cultivadas, el rodaballo posee una de las menores cantidades de ADN por célula haploide 0,65-0,87 pg (Cuñado *et al.* 2002; Hardie y Hebert, 2004) y se estima que su tamaño genómico puede estar en torno a las 800 Mb. Dado el valor de longitud genómica estimado (Ge), estos datos determinan una relación entre longitud física y genética en el genoma de rodaballo de 696,5 kb por cM.

Utilizando los tamaños genómicos recogidos en www.genomesonline.org y las correspondientes longitudes de los mapas genéticos publicados en varias especies de peces, se pueden establecer comparaciones promedio entre distintas especies. Se observa entonces un amplio rango de valores, desde 3000 kb por cM en el salmón atlántico (Moen *et al.*, 2004) o 723 kb/cM en el pez cebrá (Knapik *et al.*, 1998), a tan sólo 329,6 kb/cM en el mapa femenino del pez globo (*Takifugu rubripes*, Kai *et al.*, 2005). Dada esta variabilidad, no parece posible establecer una relación lineal entre frecuencia de recombinación y cantidad de ADN, sin embargo, se ha descrito una tendencia al aumento de la frecuencia de quiasmas por longitud del ADN al disminuir el tamaño genómico (Mb) de las especies estudiadas (Ress y Durrant, 1986; Gerber y Rodolphe, 1994). En este sentido en rodaballo se ha observado el menor grado de compactación de la cromatina descrita en peces (Cuñado *et al.*, 2002), que a su vez se ha relacionado con una mayor frecuencia de entrecruzamiento por unidad física de mapa (Cuñado *et al.*, 2002).

Este fenómeno se hace evidente en el caso del pez globo, cuyo mapa genético presenta una longitud muy similar al de especies que doblan o triplican su contenido de ADN por cM, como es el caso de medaka (795,3 kb/cM; Kimura *et al.*, 2005) o tilapia (909 kb/cM; Kocher *et al.*, 1998). La misma tendencia, aunque con valores más moderados, se aprecia en rodaballo, cuyo mapa genético preliminar presenta una longitud muy próxima al de medaka o

tilapia, a pesar de que la relación entre tamaño genético y físico en este organismo de genoma compacto es inferior a la de las especies citadas.

La similitud de las distancias génicas entre especies con importantes diferencias en cuanto al tamaño físico de sus genomas llevó a suponer un elevado nivel de conservación en el mecanismo de entrecruzamiento (Plomion *et al.*, 1995), sin embargo poco se sabe de cómo se regula la tasa de recombinación en los eucariotas superiores.

Se ha demostrado que dicha tasa varía no sólo entre especies, sino entre familias e individuos dentro de una misma especie y que resulta afectada por el sexo, la secuencia de ADN y la estructura cromosómica (Robinson, 1996), lo que se traduce en diferencias en la frecuencia de recombinación entre y dentro de los grupos de ligamiento de un mismo mapa genético (Nachman y Churchill, 1996; Ohara *et al.*, 2005). Todo ello complica la traducción de las distancias genéticas en distancias físicas y la aplicación de los mapas para la el rastreo de genes de interés y su eventual clonación posicional.

Diferencias entre sexos en las tasas de recombinación. Uno de los factores que debe considerarse a la hora de interpretar las distancias del presente estudio, es la diferencia entre individuos en la tasa de recombinación, particularmente la que se refiere a individuos de diferente sexo. Tal diferencia, si bien generalmente no va a suponer alteraciones del orden de marcadores a lo largo del mapa, sí afectará a las estimas de las distancias genéticas entre loci (Gates *et al.* 1999; Gilbey *et al.*, 2004).

En la mayoría de las especies de Vertebrados se han descrito distintas frecuencias de recombinación entre machos y hembras, siendo generalmente el sexo heterogamético el que presenta una menor tasa de recombinación. Este fenómeno se conoce como regla de Haldane (1922). El caso más extremo es el de *Drosophila*, donde la recombinación esta completamente ausente en

los machos. En mamíferos, los mapas femeninos son, generalmente, entre 1 y 2 veces más largos que los masculinos (Donis-Keller *et al.*, 1987; Jacob *et al.*, 1995; Collins *et al.*, 1996; Dietrich *et al.*, 1996).

En peces, no siempre se ha podido identificar el sexo heterogamético, ya que en muchas especies, entre las que se incluye el rodaballo, no se conoce la base genética de la determinación del sexo ni se han detectado heteromorfismos cromosómicos asociados al sexo. Sin embargo, se ha observado que los machos presentan típicamente tasas de recombinación inferiores a las hembras, variando las razones de recombinación H:M desde 1,6:1 en el pez cebra (Singer *et al.*, 2002) hasta 3,92:1 en el salmón atlántico (Gilbey *et al.*, 2004), siendo esta última una de las mayores diferencia sexo-específica conocida en Vertebrados. Dado que este fenómeno se ha manifestado en la mayoría de las especies para las que se dispone de mapas específicos de sexo (Knapik *et al.*, 1998; Sakamoto *et al.*, 2000; Moen *et al.*, 2004; Kai *et al.*, 2005), se ha especulado que esta característica podría ser general en teleósteos (Kai *et al.*, 2005).

Suponiendo que el rodaballo siga la tendencia general con una mayor tasa de recombinación en las hembras, la longitud genética del presente mapa sería posiblemente una sobreestimada de la real. Para obtener una estima más precisa de las distancias, sería interesante construir un mapa genético masculino (ya sea mediante una familia de referencia androgenética o una F₂ en la que pueda seguirse la segregación vía macho) que combinado con el actual mapa ginogenético permitiera establecer un mapa promedio para la especie (Poompuang y Na-Nakorn, 2004).

La naturaleza molecular de la supresión de la recombinación no está clara. Lindahal (1991) sugirió varias explicaciones, como el tiempo empleado en la profase meiótica en un sexo, la actividad transcripcional de ciertos genes durante la meiosis de uno de los sexos, o la presencia de secuencias reconocidas por enzimas sexo-específicas. En todo caso, la existencia de tales

diferencias tiene implicaciones prácticas. Por ejemplo, en la aplicación a la selección asistida por marcadores (Glazier *et al.*, 2002; Kai *et al.*, 2005), se espera que el sexo con una menor tasa de recombinación transmita asociaciones marcador-QTL en un ligamiento más estrecho. En otras aplicaciones, como el clonaje posicional, donde el mapeo fino es crítico, es preferible utilizar meiosis femeninas, dado que la tasa de recombinación alta ayuda a distinguir loci estrechamente ligados con un menor número de individuos (Postlethwait *et al.*, 1999; Singer *et al.*, 2002).

Diferencias en las tasas de recombinación entre y dentro de LGs. En los mapas más densos se ha encontrado que la relación Kb/cM no es constante a lo largo del mapa, sino que varía dependiendo del LG e incluso dentro de cada LG (Ohara *et al.*, 2005), relacionándose con diferencias en las tasas recombinacionales entre cromosomas y dentro de cada cromosoma. La densidad del presente mapa es demasiado escasa para establecer una comparación entre las tasas de recombinación de los diferentes grupos de ligamiento y menos aún para aventurar alguna correspondencia entre estos grupos y los cromosomas de rodaballo. Si bien la amplia variación de longitudes genéticas entre los diferentes LGs del mapa (183,0 a 7,6 cM), coincide con la diversidad de tamaños cromosómicos del cariotipo de rodaballo, el aumento en la densidad de marcadores podría redefinir las distancias de mapa y revelar nuevas asociaciones entre los grupos ya identificados y con los marcadores no ligados.

Sólo en los LGs más densos (LG1 y LG3) parecen revelarse diferencias en las tasas de recombinación dentro del grupo. Ambos LGs presentan agrupaciones de marcadores AFLP en posiciones “acrocéntricas”. Como ya se discutió en apartados previos, atendiendo a los datos disponibles no se puede resolver si la baja frecuencia de recombinación entre los marcadores implicados se debe a su localización en regiones físicamente próximas en el cromosoma o a regiones extensas donde raramente ocurre entrecruzamiento,

potencialmente heterocromatina centromérica o pericentromérica (Sakamoto *et al.*, 2000; Singer *et al.*, 2002). Si efectivamente estos loci están marcando los centrómeros podría averiguarse utilizando alguno de estos marcadores como sonda (previo aislamiento, clonación y secuenciación de la banda AFLP) y comprobar mediante hibridación *in situ* su localización cromosómica (ten Hoopen *et al.*, 1996; Strommer *et al.*, 2002).

V.CONCLUSIONES

- Las familias de ginogenéticos haploides desarrolladas en rodaballo constituyeron un material de referencia adecuado para el desarrollo del mapa genético de esta especie. Se recomienda especialmente su utilización cuando no se disponga de líneas seleccionadas consanguíneas y se trabaje con marcadores dominantes. Sin embargo, se debe poner especial atención en la selección del ADN de trabajo al tratarse de material haploide de origen embrionario. Las dificultades asociadas a la extracción de ADN de alto peso molecular a partir de embriones haploides con pocas horas de desarrollo, disminuyó el rendimiento de la tecnología AFLP, especialmente sensible a la calidad del ADN de partida. Un porcentaje importante de las bandas AFLP detectadas a partir del ADN de estos embriones, tuvieron que ser descartadas para el análisis genotípico al no alcanzar los parámetros de calidad establecidos.

- El protocolo de amplificación genómica GenomiPhi™ ayudó a resolver los problemas ligados a la escasez de ADN en los embriones haploides, proporcionando un reservorio de ADN genómico para futuras ampliaciones del mapa, sin necesidad de recurrir a estrategias como el desarrollo de líneas celulares. Este ADN amplificado resultó adecuado para el desarrollo de marcadores específicos de secuencia, sin embargo, demostró ciertos inconvenientes para la caracterización de marcadores anónimos como los AFLP. La amplificación previa del ADN molde generó fragmentos extra en los perfiles AFLPs, no presentes en los patrones resultantes del ADN original.

- La amplificación cruzada en rodaballo con marcadores microsatélite caracterizados en otras especies de peces planos resultó muy poco eficaz, lo que apunta a una escasa conservación de las regiones adyacentes de estos loci en el orden Pleuronectiformes. Estos datos desaconsejan el uso de la amplificación cruzada para la identificación de secuencias SSR en rodaballo. Alternativamente, de cara a la construcción de un mapa de segunda generación en esta especie, deberían centrarse los esfuerzos en el desarrollo

de librerías enriquecidas en secuencias repetitivas. Aunque los microsatélites son marcadores costosos, su nivel de polimorfismo, fiabilidad técnica y herencia codominante hace de ellos marcadores idóneos para el mapeo. Estos marcadores son excelentes puntos de anclaje, tanto para mapas de una misma especie originados a partir de diferentes familias de referencia o entre mapas de diferentes especies, por lo que se recomienda su uso incluso en los proyectos de mapeo de menor densidad.

- Los resultados obtenidos en esta memoria son la primera experiencia de la aplicación de la tecnología AFLP en rodaballo. Dicha tecnología resultó muy eficaz en cuanto a la obtención de marcadores en esta especie en relación al coste económico y el tiempo invertido en su desarrollo. La combinación enzimática utilizada, *EcoRI/TaqI*, rindió patrones de bandas no excesivamente saturados y con una amplia distribución de tamaños, lo que facilitó la caracterización de fragmentos. En cuanto a la productividad de los cebadores, aquellos con bases selectivas ricas en GC, rindieron un mayor número de bandas. Sin embargo, no pudo establecerse una relación entre la composición de bases de los cebadores y el nivel de polimorfismos detectados.

- Los fenómenos de amplificación inespecífica y homoplasia de tamaño, parecen tener una baja incidencia entre los fragmentos AFLP y generalmente, son ignorados en los trabajos basados en este tipo de marcadores. Sin embargo, ambos fenómenos podrían influir en la interpretación de las relaciones de ligamiento entre los marcadores multiloci. Por otro lado, la homoplasia de tamaño puede llevar a enmascarar el porcentaje real de polimorfismos AFLP. Si bien en esta memoria, sólo se comprobó el efecto de la homoplasia de tamaño en una pareja de primers AFLP, los resultados obtenidos sugieren estudiar en mayor profundidad la influencia de este fenómeno en las desviaciones de las razones de segregación mendelianas.

- El mapa genético de rodaballo recogido en esta memoria incluye 23 grupos de ligamiento formados por 118 marcadores. La presencia de un grupo

de ligamiento extra en relación al número de cromosomas de la especie, así como de 18 marcadores no ligados, indican que se trata de un mapa parcial, en el que quedan sin cubrir amplias regiones del genoma. A fin de aumentar la potencia estadística en la detección de ligamiento, sería necesario un incremento del número de marcadores, más que del número de descendientes analizados. Un número mayor de marcadores contribuiría a localizar puntos intermedios entre los loci ligados y los marcadores no ligados y a detectar posibles relaciones de ligamiento entre los grupos menores del presente mapa, lo que permitiría ajustar el número de grupos de ligamiento al número de cromosomas de la especie. Por otro lado, un aumento en el número de descendientes analizados aumentaría la precisión en el cálculo de las frecuencias de recombinación y por tanto, permitiría estimar las posiciones relativas de los marcadores con una mayor confianza estadística.

- La distribución aleatoria de los marcadores entre los grupos de ligamiento, apoya la idoneidad de los polimorfismos seleccionados (tanto microsatélite como AFLP) para el análisis de mapeo. Sólo en dos de los LGs, la disposición agrupada de algunos marcadores AFLP supera la esperada por azar. En mapas más saturados, estas agrupaciones se han identificado con regiones pericentroméricas, en el caso de rodaballo, sería necesario disponer de una mayor densidad de marcadores para establecer tal relación.

- A partir del mapa presentado en esta memoria se obtuvo una primera estima del tamaño genético del genoma femenino de rodaballo. Este dato es fundamental para planificar futuros estudios de mapeo en esta especie, ya que podrá valorarse con mayor precisión el número de marcadores necesarios para cubrir su genoma con una determinada densidad media de marcadores.

VI. BIBLIOGRAFÍA

- Ajmone-Marsan P., Negrini R., Crepaldi P., Milanese E., Gorni C. Valentini A. y Cicogna M. 2001. Assessing genetic diversity in Italian goat populations using AFLP markers. *Animal Genetics* 32, 281-288.
- Ajmone-Marsan P., Valentini A., Cassandro M., Vecchiotti-Antaldi G., Bretoni G. y Kuiper M. 1997. AFLP markers for DNA fingerprinting in cattle. *Anim. Genet.*, 28: 418-426.
- Alonso-Blanco C., Peeters J.M., Koornneef M., Lister C., Dean C., *et al.*, 1998. Development of an AFLP based linkage map of *Ler*, *Col* and *Cvi* *Arabidopsis thaliana* ecotypes and construction of a *Ler/Cvi* recombination inbred line population. *Plant J.* 14: 259-271.
- Angers B. y Bernatchez L. 1997. Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. *Mol. Biol. Evol.* 14: 230-238.
- Arai K. 1999. Genetic improvement of aquaculture finfish species by chromosome manipulation techniques in Japan. *Workshop on Biotechnology for Improvement Reproduction in Finfish Aquaculture*. Honolulu, University of Hawaii, pp.25.
- Arnheim N., Li H. y Cui X. 1990. PCR analysis of DNA sequences in single cells: single sperm mapping and genetic disease diagnosis. *Genomics* 8: 415-419.
- Baranski M., Rourke M., Loughnan S., Austin C. y Robinson N. 2006. Isolation and characterization of 125 microsatellite DNA markers in the blacklip abalone, *Haliotis rubra*. *Mol. Ecol. Notes* 6 : 740-746.
- Barnard G.A. 1949 Statistical interference. *J.R. Stat. Soc. B11*: 115-139.
- Barratt R.W., Newmeyer D., Perkins D. y Garnjobst L. 1954. Map construction in *Neurospora crassa*. *Adv. Genet.* 6:1-93.
- Belle E., Smith N. y Eyre-Walker A. 2002. Analysis of the phylogenetic distribution of isochores in vertebrates and a test of the thermal stability hypothesis. *J. Mol. Evol.* 55:356-363.
- Benfey T.J. 1999. The physiology and behavior of triploid fishes. *Rev. Fish. Sci.* 7: 39-67.
- [Benjafeld AV](#), [Wang WY](#), [Speirs HJ](#), [Morris BJ](#). 2005. Genome-wide scan for hypertension in Sydney Sibships: the GENIHUSS study. *Am J Hypertens.* 18(6):828-32.

- Bert P.F., Charmet G., Sourdille P., Hayward M.D. y Balfourier F. 1999. A high-density molecular map for ryegrass (*Lolium perenne*) using AFLP markers. *Theor. Appl. Genet.* 99:445-452.
- Bishop D.T., Cannings C., Skolnick M. y Williamson J.A. 1983. The number of polymorphic DNA clones required to map the human genome. En *Statistical Analysis of DNA Sequence Data* (Weir B.S y Dekker M., eds) New York. pp.181-200.
- Blanquer A., Alayse J.P., Berrada-Rkhami O. y Berrebi S. 1992. Allozyme variation in turbot (*Psetta maxima*) and brill (*Scophthalmus rhombus*) (Osteichthyes, Pleuronectiformes, Scophthalmidae) throughout their range in Europe. *J. Fish Biol.*, 41: 725-736.
- Bonin A., Bellemain E., Bronken Eidesen P., Pompanon F., Brochmann C. y Taberlet P. 2004. How to track and assess genotyping errors in population genetic studies. *Mol. Ecol.* 13: 3261-3273.
- Botstein D., White R.L., Skolnick M. y Davis R.W. 1980 Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314-331.
- Bouza C., Presa P., Castro J., Sánchez L. y Martínez P. 2002. Allozyme and microsatellite diversity in natural and domestic populations of turbot (*Scophthalmus maximus*) in comparison with other Pleuronectiformes. *Can. J. Fish. Aquat. Sci.* 59: 1-14.
- Bouza C., Sánchez L. y Martínez P. 1994a. Karyotypic characterization of turbot (*Scophthalmus maximus*) with conventional, fluorochrome, and restriction endonuclease banding techniques. *Mar. Biol.*, 120: 609-613.
- Bouza C., Sánchez L. y Martínez P. 1994b. Detection and interpretation of isozyme loci in turbot (*Scophthalmus maximus* L.) from population data. En: *Turbot Culture: Problems and Perspectives* (Lavens P. y Remmerswaal R.A.M., eds.). *European Aquaculture Soc.*
- Bouza C., Sánchez L. y Martínez P. 1997. Gene diversity analysis in natural populations and cultured stocks of turbot (*Scophthalmus maximus* L.). *Anim. Genet.*, 28: 28-36.
- Brooker A.L., Cook D., Bentzen P., Wright J.M. y Doyle R.W. 1994. The genomic organisation of microsatellite differs between mammals and cold water teleost fishes. *Can. J. Fish. Aquat. Sci.* 51: 1959-1966.

- Brzustowicz L.M., Merete C., Xie X., Townsend L., Gilliam T.C., *et al.* 1993. Molecular and statistical approaches to the detection and correction of errors in genotype databases. *Am. J. Hum. Genet.* 53: 1137-1145.
- Butler L. 1952. The linkage map of the tomato. *J Hered.* 43: 25-36.
- Carter T.C. y Falconer D.S. 1951. Stocks for detecting linkage in the mouse and the theory of their design. *J. Genet.* 50: 307-23.
- Castaño Sánchez C., Fuji K., Kobayashi K., Hasegawa O., Ohara E., Ozaki A., M. Coimbra M.R., Sakamoto T. y Okamoto N. 2006. Isolation and characterization of new microsatellite markers from the Japanese flounder (*Paralichthys olivaceus*). *Mol. Ecol. Notes* 6: 456-458.
- Castiglioni P., Ajmone-Marsan P., van Wijk R. y Motto M. 1999. AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. *Theor. Appl. Genet.* 99: 425-431.
- Castro J., Bouza C., Presa P., Pino-Querido A., Riaza A., Ferreiro I., Sánchez L. y Martínez P. 2004. Potential sources of error in parentage assessment of turbot (*Scophthalmus maximus*) using microsatellite loci. *Aquaculture*, 242: 119-135.
- Castro J., Bouza C., Sánchez L., Cal R.M, Piferrer F. y Martínez P. 2003. Gynogenesis assesment using microsatellite geneticc markers in turbot (*Scophthalmus maximus*). *Mar. Biotechnol.* 5: 584-592.
- Castro J., Pino A., Hermida M., Bouza C., Riaza A., Ferreiro I., Sánchez L. y Martínez P. 2006. A microsatellite marker tool for parentage analysis in Senegal sole (*Solea senegalensis*): Genotyping errors, null alleles and conformance to theoretical assumptions. *Aquaculture* 261: 1194-1203.
- Cervera M.T., Storme V., Ivens B., Gusmao J., Liu B.H. *et al.* 2001. Dense genetic linkage map of three Populus species (*P. deltoides*, *P.nigra* and *P.trichocarpa*) based on AFLP and microsatellite markers. *Genetics* 158: 787-809.
- Chagné D., Lalanne C., Madur D., Kumar S., Frigério J.M., *et al.* 2002. A high density genetic map of maritime pine based on AFLPs. *Ann. For. Sci.* 59:627-636.
- Chakravarti A., Lasher L.K. y Reefer J.E. 1991. A maximum likelihood method for estimating genome length using genetic linkage data. *Genetics* 128: 175-182.

- Chanet B. 2003. Interrelationships of Scophthalmid fishes (Pleuronectiformes: Scophthalmidae). *Cybium* 27: 275-286.
- Chistiakov D.A., Hellemans B., Tsigenopoulos C. S., Law A.S., Bartley N., Bertotto D., Libertini A., Kotoulas G., Haley C.S., y Volckaert F.A.M. 2004. Development and linkage relationships for new microsatellite markers of the sea bass (*Dicentrarchus labrax* L.). *Animal Genetics*. 35: 53-57.
- Chistiakov D.A., Hellemans B. y Volckaert F.A.M. 2005 Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. *Aquaculture* 255: 1-29.
- Clark A.G.y Lanigan C.M.S. 1993. Prospects for estimating nucleotide divergence with RAPDs. *Molecular Biology and Evolution* 10: 1096-1111.
- Clark M.S. 2003. Genomics and mapping in Teleostei (bony fish). *Comp. Funct. Genom* 4: 182-193.
- Cloutier S., Cappadocia M. y Landry S. 1997. Analysis of RFLP mapping inaccuracy in *Brassica napus* L. *Theor. Appl. Genet.* 95: 83-91.
- Cnaani A., Hallerman E.M., Ron M., Weller J.I., Indelman M., Kashi Y., Gall G.A.E. y Hulata G. 2003. Detection of a chromosomal region with two quantitative trait loci, affecting cold tolerance and fish size in an F-2 tilapia hybrid. *Aquaculture* 223: 117.
- [Cnaani A.](#), [Lee B.Y.](#), [Ozouf-Costaz C.](#), [Bonillo C.](#), [Baroiller J.F.](#), [D'Cotta H.](#), [Kocher T.](#) 2007. Mapping of sox2 and sox14 in tilapia (*Oreochromis spp.*). [Sex Dev.](#) 1(3):207-10.
- Coimbra M.R.M., Kobayashi K., Koretsugu S., Hasegawa O., Ohara E., Ozaki A., Sakamoto T., Naruse K. y Okamoto N. 2003. A genetic linkage map of the Japanese flounder, *Paralichthys olivaceus*. *Aquaculture*, 220: 203-218.
- Collins A., Teague J., Keats B.J.B. y Morton N.E. 1996. Linkage map integration. *Genomics* 36: 157-162
- [Conneally P.M.](#), [Edwards J.H.](#), [Kidd K.K.](#), [Lalouel J.M.](#), [Morton N.E.](#), [Ott J.](#), [White R.](#) 1985. Report of the Committee on Methods of Linkage Analysis and Reporting. [Cytogenet. Cell. Genet.](#) 40(1-4):356-9.
- Coughlan J., McCarthy E., McGregor D., O'Donoghue P., Galvin P., Fitzgeral R., Daemen E., Imsland A., Stefansson S. y Cross T. 1996. Four

- polymorphic microsatellites in turbot *Scophthalmus maximus*. *Anim. Genet.* 27: 433-442.
- Coughlan J., Stefansson M., Galvin P., Dillane E., Fitzgerald R. y Cross T.F. 2000. Isolation and characterization of 11 microsatellite loci in Atlantic halibut (*Hippoglossus hippoglossus* L.). *Molecular Ecology* 9: 817-829.
- Crawford A. M. Dodds K. G., Ede A. J., Pierson C. A., Montgomery G. W. *et al.* 1995. An autosomal genetic linkage map of the sheep genome. *Nature* 380: 149-152.
- Cristescu M.E.A., Colbourne J.K., Radivojc J. y Lynch. 2006. A microsatellite-based genetic linkage map of the waterflea, *Daphnia pulex*: On the prospect of crustacean genomics. *Genomics* 88: 415-430.
- Cuñado N., Terrones J., Sánchez L., Martínez P. y Santos J.L. 2001. Synaptonemal complex analysis in spermatocytes and oocytes of turbot, *Scophthalmus maximus* (Pisces, Scophthalmidae). *Genome*, 44: 1143-1147.
- Cuñado N., Terrones J., Sánchez L., Martínez P. y Santos J.L. 2002. Sex-dependent synaptic behaviour in triploid turbot, *Scophthalmus maximus* (Pisces, Scophthalmidae). *Heredity*, 89: 460-464.
- Danzmann G. y Gharbi K. 2001. Gene mapping in fishes: a means to an end. *Genetica* 111: 3-23.
- Dausset J., Cann H., Cohen D., Lathrop M., Lalouel J.M. y White R. 1990. *Centre d'Etude du Polymorphisme Humain (CEPH)*: Collaborative genetic mapping of the human genome. *Genomics* 6:575-577.
- Davey G.C., Caplice N.C., Martin S.A. y Powell R. 2001. A survey of genes in the Atlantic salmon (*Salmo salar*) as identified by expressed sequence tags. *Gene* 263: 121-130.
- Dempster A.P., Larid N.M. y Rubin D.B. 1977. Maximun likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B* 39: 1-38.
- Desoutter M.F., Chapleau T.A., Munroe B., Chanut B. y Beaunier M. 2001. Catalogue critique des types de poissons du Muséum National d'Histoire Naturelle. Ordre des *Pleuronectiformes*. *Cybio* 25 : 299-368.
- Devesa S. 1995. El Cultivo del Rodaballo (*Scophthalmus maximus*). En: *Acuicultura Mariña en Galicia*. 2as Xornadas de Medio Mariño e

- Acuicultura. Sada, Novembro-92/ Fevereiro-93. Edicións do Castro. Sada. A Coruña.
- Dib C., Faure S., Fizames C., Samson D., Drouot N., *et al.*, 1996. A comprehensive genetic map of the human genome based on 5.264 microsatellites. *Nature* 380: 149-152.
- Dietrich W., Katz H., Lincoln S.E., Shin H.S., Friedman J., Dracopoli N.C. y Lander E.S. 1992. A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* 131: 423-447.
- Dietrich W.F., Miller J., Steen R., Merchant M.A., Damron-Boles D., *et al.* 1996. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380: 152-154.
- Donaldson E.M., Devlin R.H., Piferrer F y Solar I.I. 1996. Hormones and sex control in fish with particular emphasis on salmo. *Asian Fish. Sci.* 9: 1-8.
- Donis-Keller H., Green P., Helms C., Cartinhour S., Weiffenbach B., *et al.* 1987. A genetic map of the human genome. *Cell* 51: 319-337.
- Douglas J.A., Boehnke M. y Lange K. 2000. A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. *Am. J. Hum. Genet.* 66: 1287-1297.
- Douglas J.A., Skol A.D. y Boehnke M. 2002. Probability of detection of genotyping errors and mutations as inheritance inconsistencies in nuclear-family data. *Am. J. Hum. Genet.* 70: 487-495.
- Dracopoli N.C., O'Connell P., Elsner T.I., Lalouel J.M., White R.L., Buetow K.H., Nishimura D.Y., *et al.* 1991. The CEPH consortium linkage map of human chromosome 1. *Genomics* 9: 686-700.
- Echt C.S. y Nelson C.D. 1997. Linkage mapping and genome length in eastern white pine (*Pinus strobus* L.). *Theor. Appl. Genet.* 94: 1031-1037.
- Ehm M.G., Kimmel M. y Cottingham R.W.Jr. 1996. Error detection for genetic data, using likelihood methods. *Am. J. Hum. Genet.* 58: 225-234.
- Ellis T.H.N., Turner L., Hellens R.P., Lee D., Harker C.L., Enard C., Domoney C. and Davies D.R. 1992. Linkage maps in pea. *Genetics* 130: 649-663.
- Estoup A. y Angers B. 1998. Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. In Carvalho G (Ed.). *Advances in molecular ecology* (IOS Press, Amsterdam) pp. 55-86.

- Estoup A., Gharbi K., San Cristobal M., Chevalet C., Haffray P. y Gouyomard R. 1998. Parentage assignment using microsatellites in turbot (*Scophthalmus maximus*) and rainbow trout (*Oncorhynchus mykiss*) hatchery populations. *Can. J. Fish. Aquat. Sci.* 55:715-725.
- Falconer D.S. y Mackay T.F. 1996. Introduction to Quantitative Genetics. 4ª ed. Longman, Harlow, Essex.
- Falk C.T. 1989. A simple scheme for preliminary ordering of multiple loci: application to 45 CF families. In: Elston RC, Spence MA, Hodge SE and MacCluer JW (Eds). *Multipoint Mapping and Linkage bases upon Affected Pedigree Members*. Genetic Workshop 6, Liss: New York, pp 17-22.
- Faris J.D., Laddomada B. y Gill B.S. 1998. Molecular mapping of segregation distortion loci in *Aegilops tauschii*. *Genetics* 149: 319-327.
- Felip A., Zanuy S., Carrillo M. y Piferrer F. 2001. Induction of triploidy and gynogenesis in teleost fish with emphasis on marine species. *Genetica* 111:175-195.
- Fisher R.A. 1922a. On the mathematical foundations of theoretical statistics. *Phil. Trans. Roy. Soc.* 202:309-368.
- Fisher R.A. 1922b. The systematic location of genes by means of crossover observations. *Amer. Naturalist*. 56:406-411.
- Fishman L., Kelly A.J., Morgan E. y Willis J.H. 2001. A genetic map in the *Mimulus guttatus* species complex reveals transmission ratio distortion due to heterospecific interactions. *Genetics* 159: 1701-1716.
- Fortes G.G., Bouza C., Viñas A., Martínez P. y Sánchez L. 2007. Diversity in isochore structure and chromosome Banding in fish. En: *Fish Cytogenetics* (Eds.: Pisano E., Ozouf-Costaz C., Foresti F. y Kapoor B.G.). Science Publishers, Enfield (USA). pp. 405- pp. 405-421.
- Franch R., Louro B., Tsalavouta M., Chatziplis D., Tsigenopoulus G.S., *et al.*, 2006 A genetic linkage map of the hermaphrodite teleost fish *Sparus aurata* L. *Genetics* 174: 851-861.
- Gates M.A., Kim L., Egan E.S., Cardozo T., Sirotkin H.I., *et al.* 1999. A genetic linkage map for zebrafish: comparative analysis and localization of genes and expressed sequences. *Genome Res.* 9: 334-347.
- Gerber S. y Rodolphe F. 1994. An estimation of the genome length of maritime pine (*Pinus pinaster* Ait.). *Theor. Appl. Genet.* 88: 289-292.

- Gharbi K., Gautier A., Danzmann R.G, Gharbi S., Sakamoto T. *et al.*, 2006 A linkage map for brown trout (*Salmo trutta*): chromosome homeologies and comparative genome organization with other Salmonid fish. *Genetics* 172: 2405-2419.
- Gibson R.N. 2005. *Flatfishes. Biology and Exploitation*. Blackwell Publishing. Oxford. 391pp.
- Gilbey J., Verspoor E., McLay A. y Houlihan D. 2004. A microsatellite linkage map for Atlantic salmo (*Salmo salar*). *Animal Genetics*, 35: 98-105.
- Gjedrem T. 2005 *Selection and Breeding Programs in Aquaculture*. Springer, Dordrecht.
- Glazier A.M., Nadeu J.H. y Aitman T.J. 2002. Finding genes that underline complex traits. *Science* 298: 2345-2349.
- Goldgar D.R., Fain P.R. y Kimberling W.J. 1989. Chiasma-based models of multilocus recombination: Increased power for exclusion mapping and gene ordering. *Genomics* 5:283-290.
- Gosselin I., Zhou Y., Bousquet J e Isabel N. 2002. Megagametophyte-derived linkage maps of white spruce (*Picea glauca*) based on RAPD, SCAR and ESTP markers. *Theor. Appl. Genet.* 104: 987-997.
- Grattapaglia y Sederoff. 1994. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics* 137: 1121-1137.
- Griffiths A.J.F., Miller J.H., Suzuki D.T., Lewontin R.C. y Gelbart W.M. 2002. *Genética*. (7ª ed.). McGraw-Hill-Interamericana de España. Madrid.
- Gyapay G., Morissette J., Vignal A., Dib C., Fizames C., Millasseau P., Marc S., Bernardi G., Lathrop M y Weissenbach J. 1994. The 1993-1994 Genethon human genetic linkage map. *Nature Genetics* 7: 246-399.
- Ha T.T.T., Okabe M., Morishima K., Nakayama I., Nagasawa K. y Umino T. 2007. Development and multiplex PCR amplification of new microsatellite markers for the freshwater fish, ayu (*Plecoglossus altivelis*). *Mol. Ecol. Notes* 7: 635-637.
- Haanstra J.P.W., Wye C., Verbakel H., Meijer-Dekens F., van den Berg P. *et al.* 1999. An integrated high-density RFLP-AFLP map of tomato based on two *Lycopersicon esculentum* x *L. pennellii* F2 populations. *Theor. Appl. Genet.* 99: 254-271.

- Hackett C.A y Broadfoot L.B. 2003. Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. *Heredity* 90: 33-38.
- Haldane J.B.S. 1919. The mapping function. *J. Genet.*, 8: 299-309.
- Haldane J.B.S. 1922. The part played by recurrent mutation in evolution. *Am. Nat.*, 67: 5-9.
- Hamada, H., Petrino M.G. y Kakunaga T. 1982. A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proc. Natl. Acad. Sci. USA*, 79: 6465–6469.
- Hardie, D. C., y Hebert P.D.N. 2004. Genome-size evolution in fishes. *Can. J. Fish. Aquat. Sci.* 61: 1636–1646. [
- Hartl D.L. y Jones E.W. 1998. *Genetics: Principles and Analysis*. 4ª Ed. Gareth Stevens Publishing. Sudbury, MA.
- Hartl D.L y Jones E.W. 2002. *Essential Genetics: A Genomics Perspective*. 3ª ed. Jones and Bartlett Publishers. Sudbury, MA.
- [He C.](#), [Chen L.](#), [Simmons M.](#), [Li P.](#), [Kim S.](#), y [Liu Z.J.](#) 2003. Putative SNP discovery in interspecific hybrids of catfish by comparative EST analysis. *Anim Genet.* 34(6):445-8.
- Hinegardner R. 1968. Evolution of cellular DNA content in teleost fishes. *Am. Nature* 102: 517-523.
- Hinegardner R. y Rosen D.E. 1972. Cellular DNA content and the evolution of teleostean fishes. *Am. Nature* 106: 621-644.
- Hoarau G., Cook D., Stam W.T. y Olsen J.L. 2002. New microsatellites primers for plaice, *Pleuronectes platessa* L. (Teleostei: Pleuronectidae). *Mol. Ecol. Notes*, 2: 60-61.
- Huber S. y Hedgecock D. 2004. Linkage maps of microsatellite DNA markers for the *Pacific Oyster*. *Genetics* 168: 351-362.
- Hulbert S.H., Ilott T.W., Legg E.J., Lincoln S.E., Lander E.S. y Michelmore R.W. 1988. Genetic analysis of the fungus, *Bremia lactucae* using restriction fragment length polymorphisms. *Genetics* 120: 947-958.
- Hurme P. y Savolainen O. 1999. Comparison of homology and linkage of random amplified polymorphic DNA (RAPD) markers between individual trees of Scots pine (*Pinus sylvestris* L.). *Mol. Ecol.* 8: 15-22.

- Husebye H., Lund S., Möller m., Sunde A. y Krokan H.E. 1994. A Bkm-related DNA sequence gives individual DNA fingerprints in turbot (*Scophthalmus maximus*), but neither Bkm-related, hyman SRY or human ZFY probes detect genetic sex differences. *Comp. Biochem. Physiol.* 107B: 69-73.
- Ihssen P.E., McKay L.R., McMillan I. y Phillips R.B. 1990. Ploidy manipulation and gynogenesis in fishes: cytogenetic and fisheries applications. *Trans Am Fish Soc.* 119:698-717.
- Imsland A.K., Folkvord A., Grung G.L. Stefansson S.O. y Taranger G.L. 1997. Sexual dimorphism in growth and maturation of turbot, *Scophthalmus maximus*. *Aquac. Res.* 28:101-114.
- Iyengar A., Piyapattanakorn S., Heipel D.A., Stone D.M., Howell B.R., Child A.R. y Maclean N. 2000. A suitable polymorphic microsatellite markers in turbot (*Scophthalmus maximus*) with potential for use across several flatfish species. *Molecular Ecology* 9: 365-378
- Jacob H.J., Brown D.M., Bunker R.K., Daly M.J., Dzau V.J. *et al.* 1995. A genetic map of the laboratory rat, *Rattus norvegicus*. *Nat. Genet.* 9: 63-69.
- Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N, *et al.*, 2004 Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431: 946-957.
- Jensen J. y Jorgensen J.H. 1975a. The barley chromosome 5 linkage map. *Hereditas* 80: 5-16.
- Jensen J. y Jorgensen J.H. 1975b. The barley chromosome 5 linkage map II. Extension of the map with four loci. *Hereditas* 80: 17-26.
- Jilin L. 2003. The development direction of turbot farming Industry in China. *Scientific Fish Farming* 7: 26-28.
- Jin H., Domier L.L., Kolb F.L., Brown C.M. 1998. Identification of quantitative loci for tolerance to barley yellow dwarf virus in oat. *Phytopathology* 88: 410-415.
- Johnson S.L., Gates M.A., Johnson M., Talbot W.S., Horne S., Baik K., Rude S., Wong J.R. y Postlewait J.H. 1996. Centromere-linkage analysis and consolidation of the zebrafish genetic map. *Genetics* 142: 1277-1288.
- Jones A. 1974. Sexual maturity, fecundity and growth of the turbot (*Scophthalmus maximus*). *J. Biol. Ass. UK.* 54: 109-125.

- Kai W., K. Kikuchi, M. Fujita, H. Suetake, A. Fujiwara *et al.*, 2005 A genetic linkage map for the tiger pufferfish, *Takifugu rubripes*. *Genetics* 171: 227-238.
- Karlin S. 1984. Theoretical aspects of genetic map functions in recombination processes. In: *Human population genetics: The Pittsburgh symposium*, edited by A Chakravarti, 209-228. Van Nostrand Reinhold, New York.
- [Kasahara M.](#), [Naruse K.](#), [Sasaki S.](#), [Nakatani Y.](#), [Qu W.](#), *et al.* 2007. The medaka draft genome and insights into vertebrate genome evolution. [Nature](#). 447(7145):714-719.
- Keim P.J., Schupp J.M, Travis S.E., Clayton K., Zhu T. *et al.* 1997. A high-density soybean genetic map based on AFLP markers. *Crop. Sci.* 37: 537-543.
- Kesseli R.V., Paran I. y Michelmore R.W. 1994. Analysis of a detailed genetic linkage map of *Lactuca sativa* (lettuce) constructed from RFLP and RAPD markers. *Genetics* 136: 1435-1446.
- Khoo G., Lim M.H., Suresh H., Gan D.K.Y., Lim K.F., Chen F., Chan W., Lim T.M y Phang P.E. 2003. Genetic linkage maps of the guppy (*Poecilia reticulata*): Assignment of RAPD markers to multipoint linkage groups. *Marine Biotechnology*. 5: 279-293.
- Kijas J.M.H., Fowler J.C.S., Garbett C.A. y Thomas M.R. 1994. Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *Bio-Techniques* 16: 656-662.
- Kimura T., Yshida K., Shimada A., Jindo T., Sakaizumi M., Mitani H., Naruse K., Takeda H., Inoko H., Tamiya G. y Shinya M. 2005. Genetic linkage map of medaka with polymerase chain reaction length polymorphisms. *Gene* 363: 24-31.
- Knapik E.W., Goodman A., Ekker M., Chevrette M., Delgado J., *et al.* 1998. A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nature Genetics* 18: 338-343.
- Knapp M., Neugebauer M., Fimmers R., Seuchter S.A., Baur M.P. 1989. Preliminary ordering of multipoint linkage data. In: Elston RC, Spence MA, Hodge SE, MacCluer JW (eds) Multipoint mapping and linkage based upon affected pedigree members. Genetic Analysis Workshop 6. Liss, New York, pp 41-46

- Kocher T. D., Lee W. J., Sobolewska H., Penman D. y Mcandrew B. 1998. A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). *Genetics* 148: 1225-1232.
- Kondo Y., Mori M., Kuramoto T., Yamada J., Beckmann J.S., Simon-Chazottes D., Montagutelli X., Guénet J.L., Serikawa T. 1993. DNA segments mapped by reciprocal use of microsatellite primers between mouse and rat. *Mamm Genome* 4(10): 571-6.
- Kosambi D. D. 1944 The estimation of map distances from recombination values. *Ann. Eugenics* 12: 172-175.
- Kruglyak L., Daly M.J., Reeve-Daly M.P., Lander E.S. 1996. Parametric and nonparametric linkage analysis: a unified multipoint approach *Am. J. Hum. Genet.* 58:1347-1363.
- Kruglyak L., Daly M.J. y Lander E.S. 1995. Rapid multipoint linkage analysis of recessive traits in nuclear families, including homozygosity mapping. *Am. J. Hum. Genet.* 56: 519-527.
- Kuang H, Richardson T., Carson S., Wilcox P. y Bongarten B. 1999. Genetic analysis of inbreeding depression in plus tree 850.55 of *Pinus radiata* D. Don. I. Genetic map with distorted markers. *Theor. Appl. Genet.* 98: 697-703.
- Lander E.S. y Bostein D. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121: 185-199.
- Lander E.S. y Green P. 1987. Construction of multilocus linkage maps in humans. *Proc Natl Acad Sci USA* 84: 2363-2367.
- Lander E.S., Green P., Abrahamson J., Barlow A., Daly M.J., Lincoln S. y Newburg L. 1987. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1: 174-181.
- Lange K. y Boehnke M. 1982. How many polymorphic genes will it take to span the human genome. *Am. J. Hum. Genet.* 34: 842-845.
- Lathrop G.M. y Lalouel J.M. 1984. Easy calculation of lod scores and genetic risks on small computers. *Am. J. Hum. Genet.* 36: 460-465.
- Launey S., y Hedgecock D. 2001. High genetic load in the Pacific oyster *Crassostrea gigas*. *Genetics* 159: 255-265.

- Lee B.Y., Lee W.L., Streelman J.T., Carleton K.L. y Aimee E.H. 2005. A second-generation genetic linkage map of tilapia (*Oreochromis spp.*). *Genetics* 170: 237-244.
- Lee M.K., Blake S.L., Singer J.T. y Nicholson B.L. 1996 Genomic variation of aquatic birnaviruses analyzed with restriction fragment length polymorphisms. *Appl. Environ. Microbiol.* 62: 2513-2520.
- Lewin H.A., Wheeler M.B. y Schook L.B. 1990 UFO-PCR: A strategy for linkage mapping in species with underdeveloped genetic maps. In: J.E. Womack (Ed.). *Mapping the Genomes of Agriculturally Important Animals*. Pp. 95-102. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Li H., Gyllenstein V.B., Cui X., Saiki R.K., Erlich H.A. y Arnheim N. 1988. Amplification and analysis of DNA sequences in single human sperm. *Nature* 335: 414-417.
- Li L. y Guo X. 2004. AFLP-based genetic linkage map of the pacific oyster *Crassostrea gigas* thunberg. *Mar. Biotechnol.*, 6: 26-36.
- Li L., Xiang J., Liu X., Zhang Y., Dong B. y Zhang X. 2005. Construction of AFLP-based genetic linkage map for Zhikong scallop, *Chlamys farreri* Jones et Preston and mapping of sex-linked markers. *Aquaculture* 245: 63-73.
- Lidner K.R., Seeb J.E., Habicht C., Knudsen K.L., Kretschmer E., Reedy D.J., Spruell P. y Allendorf F.W. 2000. Gene-centromere mapping of 312 loci in pink salmon by half-tetrad analysis. *Genome* 43: 538-549.
- Lie O., Slettan A., Lingaas F., Olsaker I., Hordvik I. et al., 1994 Haploid gynogenesis: a powerful strategy for linkage analysis in fish. *Anim. Biotechnol.* 5: 33-45.
- Lincoln S.E. y Lander E.S. 1992. Systematic detection of errors in genetic linkage data. *Genomics* 14: 604-610.
- Lindahl K.F. 1991. His and hers recombinational hotspots. *Trends Genet.* 7: 273-276.
- Litt M. y Luty J.A. 1989. A hypervariable microsatellite revealed by in vitro amplification of dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* 44, 397-401.
- Liu B.H. 1998. *Statistical Genomics: Linkage, mapping and QTL analysis*. CRC Press.

- Liu Q., Goudie C.A., Simco B.A., Davis K.B. y Morizot D.C. 1992. Gene-centromere mapping of six enzyme loci in gynogenetic channel catfish. *J. Hered.*, 83: 245-248.
- Liu Z., Karsi A. y Dunham R.A. 1999. Development of polymorphic EST markers suitable for genetic linkage mapping of catfish. *Mar. Biotechnol.*, 1: 437-447.
- Liu Z., Karsi A., Li P., Cao D. y Dunham R. 2003. An AFLP-based genetic linkage map of Channel Catfish (*Ictalurus punctatus*) constructed by using an interspecific hybrid resource family. *Genetics* 165: 687-694.
- Liu Z.J. y Cordes J.F. 2004. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture*, 238: 1-37.
- Lynch M. y Walsh B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates Inc, Sunderland, Massachusetts, USA
- Maliepaard C., Jansen J. y Van Ooijen J.W. 1997. Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences for applications. *Genet. Res., Camb.* 70: 237-250.
- Manly K.F., Cudmore Jr. y Meer J.M. 2001. Map Manager QTX, cross-platform software for genetic mapping. *Mammalian Genome*. 12: 930-932.
- Martínez P. 2005. Aplicaciones de la genética para la mejora de la acuicultura. *Bol. Inst. Esp. Oceanogr.* 21: 225-238
- McConnell S.K., Beynon C., Leamon J. y Skibinski D.O. 2000. Microsatellite marker-based genetic linkage maps of *Oreochromis aureus* and *O. Niloticus* (Cichlidae); extensive linkage group segment homologies revealed. *Anim. Genet.* 31: 214-218.
- Mikawa S., Akita T., Hisamatsu N., Inage Y., Ito Y., *et al.* 1999. A linkage map of 243 markers in an intercross of *Gottingen miniature* and *Meishan pigs*. *Anim. Genet.* 30: 407-417.
- Moen T., Hoyheim B., Munck H. y Gomez-Raya L. 2004. A linkage map of Atlantic salmon (*Salmo salar*) reveals an uncommonly large difference in recombination between the sexes. *Anim. Genet.* 35: 81-92.
- Morgan T.H. 1928. *The theory of genes*. Yale University Press, New Haven, Conn.

- Morton N. 1955a. Sequential test for the detection of linkage. *Am. J. Hum. Genet.* 7: 277-318.
- Morton N.E. 1955b LODs past and present. *Genetics* 140: 7-12.
- Muller H.J. 1916. The mechanism of crossing over. *Am. Nat.* 50: 193-207.
- Muller U.G. y Wolfenbarger L. 1999. AFLP genotyping and fingerprinting. *Tree* 14 (10): 389-394.
- Nachman M.W. y Churchill G.A. 1996. Heterogeneity in rates of recombination across the mouse genome. *Genetics* 142: 537-548.
- [Nagy I.](#), [Stágel A.](#), [Sasvári Z.](#), [Röder M.](#), [Ganal M.](#) 2007. Development, characterization, and transferability to other Solanaceae of microsatellite markers in pepper (*Capsicum annuum* L.). [Genome](#). 50(7):668-88.
- Naruse K. y Shima A. 1989. Linkage relationships of gene loci in medaka, *Oryzias latipes* (Pisces: Oryziatidae), determined by backcrosses and gynogenesis. *Biochemical Genetics* 27 (3/4): 183-198.
- Naruse K., Fukamachi S., Mitani H., Kondo M., Matsuoka T. *et al.*, 2000 A detailed linkage map of medaka, *Oryzias latipes*: comparative genomics and genome evolution. *Genetics* 157: 1773–1784.
- Neale D.B., Sewell M.M. y Brown G.R. 2002. Molecular dissection of the quantitative inheritance of wood property traits in loblolly pine. *Ann. For. Sci.* 59: 595-605.
- Negi M.S., Devic M., Delseny M. y Lakshmikumaran M. 2000. Identification of AFLP fragments linked to seed coat colour in Brassica juncea and conversion to a SCAR marker for rapid selection. *Theor. Appl. Genet.* 101: 146-152.
- Nei M. 1987. *Molecular Evolution Genetics*. Columbia University Press, New York.
- Nekrutenko A. y Li. W.H. 2000. Assessment of compositional heterogeneity within and between eukaryotic genomes. *Genome Research* 10(12):1896-1995.
- Nelson J.S. 1994. *Fishes of the world*. Wiley: New York.
- Nichols K.M., Young W.P., Danzmann R.G., Robinson B.D., Rexroad C. *et al.* 2003. A consolidated linkage map of rainbow trout (*Oncorhynchus mykiss*). *Animal Genetics*, 34: 102-115.

- Nielsen J.G. 1986. Scophthalmidae. En: *Fishes of the North-Eastern Atlantic and the Mediterranean* (Whitehead P.J.P., Bauchot M.L., Hureau J.C. y Tortonese E., eds). UNESCO, Paris, France. pp: 1287-1293.
- Nikaido A., Yoshimura H., Tsumura Y., Suyama Y., Murai M. y Nagasaka K. 1999. Segregation distortion for AFLP markers in *Cryptomeria japonica*. *Genes Genet. Syst.* 74: 55-59.
- Nuez F., 2000 Introducción al análisis del ligamiento y cartografía de marcadores simples. En: *Los Marcadores Genéticos en la Mejora Vegetal*. (Nuez F. y Carrillo J.M., Eds). Editorial de la Universidad Politécnica de Valencia. Valencia. pp: 153-293.
- Nusslein-Volhard C. 1994. Of flies and fish. *Science* 266: 572-574.
- O'Connell J.R y Weeks D.E. 1998. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am. J. Hum. Genet.* 63: 259-266.
- O'Connell J.R y Weeks D.E. 1999. An optimal algorithm for automatic genotype elimination. *Am. J. Hum. Genet.* 65: 1733-1740.
- O'Connell M. y Wright J.M. 1997. Microsatellite DNA in fishes. *Rev. Fish Biol. Fish.*, 7: 331-363.
- O'Hanlon P.C. y Peakall R. 2000. A simple method for the detection of size homoplasy among amplified fragment length polymorphism fragments. *Mol. Ecol.*, 9: 815-816.
- Ohara E., Nishimura T., Nagakura Y., Sakamoto T., Mushiake K. y Okamoto N. 2005. Genetic linkage map of two yellowtails (*Seriola quinqueradiata* and *Seriola lalandi*). *Aquaculture* 244: 41-48.
- Ohno S. 1974. Animal Cytogenetics, Vol. 4: *Chordata* 1: *Procordata, Cyclostoma and Pisces*. Gebruder Borntraeger: Berlin, Stuttgart.
- Ostrander E.A., Jong P.M., Rine J. y Duyk G. 1992. Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. *Proc. Natl. Acad. Sci. USA* 89: 3419-3423.
- Otsen M., den Bieman M., Kuiper M.T.R., Pravenec M., Kren V., Kurtz T.W., Jacob H.J., Lankhorst A. y Van Zutphen B.F.M. *et al.* 1996. Use of AFLP markers for gene mapping and QTL detection in the rat. *Genomics* 37: 289-294.

- Ott J. 1974. Estimation of the recombination fraction in human pedigrees -efficient computation of the likelihood for the human linkage studies. *Am. J. Hum. Genet.* 26: 588-597.
- Ott J. 1976. A computer program for linkage analysis of general human pedigrees. *Amer. J. Hum. Genet.* 38: 868-883.
- Ott J. 1991. *Analysis of Human Genetic Linkage*. Johns Hopkins University Press. Baltimore.
- Ott J. 1993. Detecting marker inconsistencies in human gene mapping. *Hum. Hered.* 43: 25-30.
- Ott J. 1999. *Analysis of Human Genetic Linkage*. (3^a ed) Johns Hopkins Press, Baltimore, MD.
- Ozaki A., Sakamoto T., Khoo S., Nakamura K., Coimbra M.R., Akutsu T. y Okamoto N. 2001. Quantitative trait loci (QTLs) associated with resistance/susceptibility infectious pancreatic necrosis virus (IPNV) in rainbow trout (*Oncorhynchus mykiss*). *Mol. Genet. Genom.* 265: 25-31.
- Paglia G.P., Olivieri A.M. y Morgante M. 1998. Towards second-generation STS (sequence-tagged sites) linkage maps in conifers: a genetic map of Norway spruce (*Picea abies* K.). *Mol. Gen. Genet.* 258: 466-478.
- Papa R., Troggio M., Ajmone Marsan P. y Nonnis Marzano F. 2005. An improved protocol for the production of AFLP markers in complex genomes by means of capillary electrophoresis. *J. Anim. Breed. Gen.*, 122: 62-68.
- Pardo B.G., Bouza C., Castro J., Martínez P. y Sánchez L. 2001. Localization of ribosomal genes in Pleuronectiformes using Ag- and CMA3 banding and in situ hybridization. *Heredity*, 86: 531-536.
- Pardo B.G., Casas L., Fortes G.G., Bouza C., Martínez P., Clark M.S. y Sánchez L. 2005. New microsatellite markers in turbot (*Scophthalmus maximus*) derived from an enriched genomic library and sequence databases. *Mol. Ecol. Notes*, 5: 62-64.
- Patterson D. 1991. Report of the Second International Workshop on Human Chromosome 21 mapping. *Cytogenet. Cell. Genet.* 57: 168-174.
- Peakall R., Gilmore S., Keys W., Morgante M. y Rafalski A. 1998. Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications

- for the transferability of SSRs plants. *Molecular Biology and Evolution* 15: 1275-1287.
- Pejic I., Ajmone-Marsan P., Morgante M., Kozumplick V., Castiglioni P., Taramino G. y Motto M. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs y AFLPs. *Theor. Appl. Genet.* 97: 1248-1255.
- Pérez, F., Erazo C., Shinaula M., Volckaert F. y Calderon J. 2004. A sex-specific linkage map of the white shrimp *Penaeus (Litopenaeus) vannamei* based on AFLP markers. *Aquaculture* 242: 105-118.
- Périn C., Hagen L.S., De Conto V., Katzir N., Danin-Poleg Y., Portnoy V., Baudracco-Arnas S., Chadouef J., Dogimont C. y Pitrat M. 2002. A reference map of *Cucumis melo* based on two recombinant inbred line populations. *Theor. Appl. Genet.*, 104: 1017-1034.
- Perkins D.D. 1962. Crossing-over and interference in a multiply marked chromosome arm of *Neurospora*. *Genetics* 47: 1253-1274.
- Piferrer F., Cal R.M, Álvarez-Blázquez B., Sánchez L. y Martínez P. 2000. Induction of triploidy in turbot (*Scophthalmus maximus*). I. Ploidy determination and the effects of cold shocks. *Aquaculture*. 188:79-79.
- Piferrer F., Cal R.M., Gómez C., Álvarez-Blázquez B., Castro J. y Martínez P. 2004. Induction of gynogenesis in turbot (*Scophthalmus maximus*): Effects of UV irradiation on sperm motility, the Herwig effect and viability during the first 6 months of age. *Aquaculture* 238: 403-419.
- Piferrer F., Cal R.M., Gómez C., Bouza C. y Martínez P. 2003. Induction of triploidy in the turbot (*Scophthalmus maximus*). II. Effects of cold shock timing and induction of triploidy in a large volume of eggs. *Aquaculture* 220: 821-831.
- Piferrer F., Felip A. y Blázquez M. 1995. Control genético y fisiológico de las proporciones de sexos de los teleósteos y su aplicación en acuicultura. En: *Aulas del Mar, Acuicultura, Biología Marina*. Zamora S., Agulleiro y García O. (eds.), Murcia. pp: 75-109.
- Pinera J.A., Bernardo D., Blanco G., Vázquez E. y Sánchez J.A. 2006. Isolation and characterization of polymorphic microsatellite markers in *Pagellus bogaraveo*, and cross-species amplification in *Sparus aurata* and *Dicentrarchus labrax*. *Mol. Ecol. Notes*, 6

- Plomion C., Bahrman N., Durel C.E. y O'Malley D.M. 1995. Genomic mapping in *Pinus pinaster* (maritime pine) using RAPD and protein markers. *Heredity* 74: 661-668.
- Poompuang S. y Na-Nakorn U. 2004. A preliminary genetic map of walking catfish (*Clarias macrocephalus*). *Aquaculture* 232: 195-203.
- Postlethwait J., Amores A., Force A. y Yan Y.L. 1999. The zebrafish genome. In: *Methods in cell biology*. Vol. 60. Academic press.
- Postlethwait J., Johnson S.L., Midson C.N., Talbot W.S., Gates M. *et al.*, 1994. A genetic linkage map for the zebrafish. *Science* 264: 699-703.
- Powell W., Morgante M., Andre C., Hanafey M., Vogel J., Tingey S. y Rafalsky A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germoplasm analysis. *Mol. Breed.* 2: 255-238.
- Powell W., Thomas B.H., Baird E., Lawrence P., Booth A., *et al.* 1997. Analysis of quantitative traits in barley by the use of amplified fragment length polymorphisms. *Heredity* 79: 48-59.
- Purdom C.E. 1976. Genetic techniques in flatfishes culture. *J. Fish. Res. Board. Can.* 33: 1088-1093.
- Purdom C.E., Jones A. y Lincoln R.F. 1972. Cultivation trials with turbot (*Scophthalmus maximus*). *Aquaculture*, 1: 213-230.
- Qi X. y Lindhout P. 1997. Development of AFLP markers in barley. *Mol. Gen. Genet.* 254: 330-336.
- Qi X., Stam P. y Lindhout P. 1998. Use of locus-specific AFLP markers to construct a high-density molecular map in barley. *Theor. Appl. Genet.* 96: 376-384.
- Quarrie S.A., Laurie D.A., Zhu J., Lebreton C., Semikhodskii A., Steed A., Witsenboer H. y Calestani C.. 1997. QTL analysis to study the association between leaf size and abscisic acid accumulation in droughted rice leaves and comparisons across cereals. *Plant. Mol. Biol.* 35: 155-165.
- Raimuzzo J.R., Reitan K.I., Jorgensen L. y Olsen Y. 1994. Lipidid-composition in turbot larvae fed live feed cultured by emulsion of different lipid classes. *Comp Biochem. Physiol. A.* 107: 699-710.

- Rajapakse S., Byrne D.H., Zhang L., Anderson N., Arumuganathan K. y Ballard R.E. 2001. Two genetic linkage maps of tetraploid roses. *Theor. Appl. Genet.*, 103: 575-583.
- Raymond M. y Rousset F. 1995. GENEPOP (Version 1.2): Population genetics Software for Exact Tests and Ecumenicism. *J. Hered.*, 86: 248-249.
- Reiter R.S., Williams J.G.K., Feldmann K.A., Rafalski J.A., Tingey S.V. y Scolnik P.A. 1992. Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. *Proc. Natl. Acad. Sci. USA* 89: 1477-1481.
- Remington D.L., Whetten R.W., Liu B.H. y O'Malley D.M. 1999. Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. *Theor. Appl. Genet.* 98: 1279-1292.
- Ress H. y Durrant A. 1986. Recombination and genome size. *Theor. Appl. Genet.* 73: 72-76.
- Rico C., Ibrahim K.M., Rico I. y Hewitt G.M. 1997. Stock composition in North Atlantic populations of whiting using microsatellite markers. *J. Fish Biol.*, 51: 462-475
- Rico C., Rico I. y Hewitt G. 1996. 470 million years of conservation of microsatellite loci among fish species. *Proc. Roy. Soc. London, Series B, Biol. Sci.*, 263: 549-557.
- Robinson B.D., Wheeler P.A., Sundin K., Sikka P. y Thorgaard G.H. 2001. Composite interval mapping reveals a major locus influencing embryonic development rate in rainbow trout (*Oncorhynchus mykiss*). *J. Hered.* 92: 16-22.
- Robinson W.P. 1996. The extent, mechanism, and consequences of genetic variation for recombination rate. *Am. J. Hum. Genet.* 59: 1175-1183.
- Rodríguez M. y Díaz O. 1990. Producción de alevines de rodaballo (*Scophthalmus maximus*). Hojas divulgadoras. Ministerio de Agricultura, Pesca e Alimentación. Secretaría general de estructuras agrarias. Nº 3.
- Roest Crollius H., Jaillon O., Bernot A., Dasilva C., Bouneau L. *et al.* 2000. Estimate of human gene number provided by genome-wide analysis using *Tetraodon nigroviridis* DNA sequence. *Nat. Genet.* 25: 235-238.

- Rothenberg E.V. 2001. Mapping of complex regulatory elements by pufferfish/zebrafish transgenesis. *Proc Natl Acad Sci USA* 98: 6540-6542.
- Roupe Van Der Voort J.N.A.M., van Eck H.J., Folkertsma R.T., van Zandvoort P.V., Hutten R.C.B., Draaistra J., Gommers F.J., Jacobsen E., Helder J. y Bakker J. 1998. Locus specificity of AFLP markers used to align genetic maps from different potato genotypes. *Mol. Gen. Genet.* 255: 438-447.
- Rozen S. y Skaletsky H.J. 2000. PRIMER3 code available at http://frodo.wi.mit.edu/primer3/primer3_code.html. In: Bioinformatics Methods and Protocols: Methods in Molecular Biology (eds Krawetz S, Misener S), pp. 365-386. Humana Press, Totowa, NJ. USA.
- Saal B. y Wricke G. 2002. Clustering of amplified length polymorphism markers in a linkage map of rye. *Plant breeding* 121: 117-123.
- Sachidanandam R., Weissman D., Schmitdt S.C., *et al.* 2001 A map of human genome sequence variation containing 1.42 million single nucleotide polymorphism. *Nature* 409: 928-933.
- Sakamoto T., Danzmann R.G., Gharbi K., Howard P., Ozaki A. *et al.*, 2000 A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. *Genetics* 155: 1331-1345.
- Sambrook J., Fritsch E.F. y Maniatis T. 1989. *Molecular Cloning: A Laboratory Manual* (2^a ed.). Cold Spring Harbor, N.Y.
- Schoen D.J. 2000. Comparative genomics, marker density and statistical analysis of chromosomal rearrangements. *Genetics* 154: 943-952.
- Sekino M. y Hara M. 2000. Isolation and characterization of microsatellite DNA loci in Japanese flounder *Paralichthys olivaceus*. *Mol. Ecol.*, 9: 2200-2202.
- Sewell M.M., Bassoni D.L., Megraw R.A., Wheeler N.C. y Neale D.B. 2000. Identification of QTL influencing wood property traits in loblolly pine (*Pinus taeda* L.). I. Physical wood properties. *Theor. Appl. Genet.* 101: 1273-1281.
- Shimoda, N., E. W. Knapik, J. Ziniti, C. Sim, E. Yamada *et al.*, 1999 Zebrafish genetic map with 2000 microsatellite markers. *Genomics* 58: 219-232.

- Singer A., Perlman H., Yan Y., Walker C., Corley-Smith G., *et al.* 2002 Sex-specific recombination rates in zebrafish (*Danio rerio*). *Genetics* 160: 649-657.
- Slasky R.J. 1985. La piscicultura europea y escandinava. *El campo* 99: 128-132.
- Slettan A., Olsaker I. y Lie O. 1997. Segregation studies and linkage analysis of Atlantic salmo microsatellites using haploid genetics. *Heredity* 78: 620-627.
- [Somers DJ](#), [Isaac P](#), [Edwards K](#). 2004. A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). [Theor Appl Genet.](#) 109(6):1105-14.
- Southern E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98: 503-517.
- Stam P y van Ooijen J.W. 1995. *JOINMAP version 2.0: software for the calculation of genetic linkage maps*. DLO-Center for Plant Breeding and Reproduction Research, Wageningen, The Netherlands.
- Stam P. 1993. Construction of integrated genetic linkage maps by means of a new computer package: JOINMAP. *Plant J.* 3: 739-744.
- Stemshorn K. C., Nolte A.W. y Tautz D. 2005 A genetic map of *Cottus gobio* (Pisces, Teleostei) based on microsatellites can be linked to the physical map of *Tetraodon nigroviridis*. *J. Evol. Biol.* 18: 1619-1624.
- Strachan T. y Read A.P. 1999. *Genética Molecular Humana*. Ediciones Omega, S.A., Barcelona.
- Streelman J.T., Zardoya R., Meyer A. y Karl S.A. 1998. Multilocus phylogeny of cichlid fishes (Pisces: Perciformes): evolutionary comparison of microsatellite and single-copy nuclear loci. *Mol. Biol. Evol.* 15: 798-808.
- Stringham H.M. y Boehnke M. 1996. Identifying marker typing incompatibilities in linkage analysis. *Am. J. Hum. Genet.* 59: 946-950.
- Strommer J., Peters J., Zethof J., de Keukeleire P. y Gerats T. 2002. AFLP maps of *Petunia hybrida*: building maps when markers cluster. *Theor. Appl. Genet.* 105:1000-1009.
- Sturtevant A.H. 1913 The linear arrangement of sex-linked factors in *Drosophila*, as shown by their mode of association. *J. Exp. Zool.*, 14: 43-59.

- Sun X.W. y Liang L.Q. 2004. A genetic linkage map of common carp (*Cyprinus carpio* L.) and mapping of a locus associated with cold tolerance. *Aquaculture* 238: 165-172.
- Tan Y.D., Wan C., Zhu Y., Lu C., Xiang Z. *et al.* 2001. An amplified fragment length polymorphism map of the silkworm. *Genetics* 157: 1277-1284.
- Tanksley S.D., Ganai M.W., Prince J.P., de Vicente M.C., Bonierbale M.W., 1992. High-density molecular linkage maps of the tomato and potato genomes. *Genetics* 132: 1141-1160.
- Tautz D. 1989. Hypervariability of simple sequences as general source for polymorphic DNA markers. *Nucleic Acids Res.*, 17: 6463-6471.
- Tautz D. y Renz M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res.*, 12: 4127-4138.
- ten Hoopen R., Robbins T.P., Fransz P.F., Montijn B.M., Oud O., Gerats A.G.M. y Nanninga N. 1996. Localization of T-DNA insertions in *Petunia* by fluorescence *in-situ* hybridization: physical evidence for suppression of recombination. *Plant Cell* 8: 823-830
- Thorgaard G.H. 1983. *Chromosome set manipulation and sex control in fish*. Fish Physiology, Vol. IXB., Hoar, W.H., Randall, D.J. and Donaldson, E.M. (Eds.). New York, N.Y.: Academic Press, pp 405-434.
- Thorgaard G.H. 1986. Ploidy manipulation and performance. *Aquaculture* 57: 57-64.
- Thorgaard G.H., Allendorf F.W. y Knudsen K.L. 1983. Gene-centromere mapping in rainbow trout: high interference over long map distances. *Genetics* 103: 771-783.
- Tiersch T.R., Chandler R.W., Wachtel S.S. y Elias S. 1989. Reference standards for flow cytometry and application in comparative studies of nuclear DNA content. *Cytometry* 10: 706-710.
- Tulseriam L.K., Glaubitz J.C., Kiss G. y Carlson J.E. 1992. Single tree genetic linkage mapping in conifers using haploid DNA from megagametophytes. *Biotechnology* (NY) 10(6): 686-690.
- Vallejos C.E., Sakiyama N.S. y Chase C. 1992. A molecular marker-based linkage map of *Phaseolus vulgaris* L. *Genetics* 131: 733-740.
- Vekemans X., Beauwens T., Lemaire M. y Roldán-Ruiz I. 2002. Data from amplified fragment length polymorphism (AFLP) markers show

- indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Mol. Ecol.* 11: 139-151.
- Verhaegen D. y Plomion C. 1996. Genetic mapping in *Eucalyptus urophylla* y *Eucalyptus grandis* using RAPD markers. *Genome* 39: 1051-1061.
- Virk P.S., Ford-Lloyd B.V. y Newbury H.J. 1998. Mapping AFLP markers associated with subspecific differentiation of *Oryza sativa* (rice) and an investigation of segregation distortion. *Heredity* 81: 613-620.
- Voorrips R.E., Jogerius M.C. y Kanne H.J. 1997. Mapping of two genes for resistance to clubroot (*Plasmodiophora brassicae*) in a population of double haploid lines of *Brassica oleracea* by means of RFLP and AFLP markers. *Theor. Appl. Genet.* 94: 75-82.
- Voorrips, R.E., 2002. MapChart: Software for the graphical presentation of linkage maps and QTLs. *The Journal of Heredity* 93 (1): 77-78.
- Vos P. y Kuiper M. 1996. *DNA markers: protocols application and overviews*. J. Wiley and Sons, Inc., Bognor Regis, Sussex.
- Vos P., Hogers R., Bleeker M., Reijans M., Van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M. y Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.*, 23: 4407-4414.
- Vuylsteke M.R., Mank R., Antonise R., Bastiaans E., Senior M.L., *et al.* 1999. Two high density AFLP linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. *Theor. Appl. Genet.* 99: 921-935.
- Wada H., Naruse K., Shimoda A. y Shima A. 1995. Genetic linkage map of a fish, the Japanese medaka *Oryzias latipes*. *Mol. Mar. Biotechnol.* 4: 269-274.
- Walter R.B., Rains J.D., Russell J.E., Guerra T.M., Daniels C., *et al.* 2004. A microsatellite genetic linkage map for xiphophorus. *Genetics* 168: 363-372.
- [Wang C.M.](#), [Lo L.C.](#), [Zhu Z.Y.](#), [Yue G.H.](#) 2006. A genome scan for quantitative trait loci affecting growth-related traits in an F1 family of Asian seabass (*Lates calcarifer*). [BMC Genomics](#). 26;7:274
- Wang C.M., Zhy Z.Y., Lo L.C., Feng F., Lin G. *et al.* 2007. A microsatellite linkage map of barramundi, *Lates calcarifer*. *Genetics* 175: 907-915.
- Watanabe T., Fujita H., Yamasaki K., Seki S. y Taniguchi N. 2004. Preliminary study on linkage mapping based on microsatellite DNA and AFLP

- markers using homozygous clonal fish in ayu (*Plecoglossus altivelis*). *Mar. Biotechnol.* 6: 327-334.
- Watson D.J., Hopkins N.H., Roberts J.W., Argetsinger Steitz J. y Weiner A.M. 1987. *Molecular Biology of the Gene*, Vol 1: General Principles, 4th edition. Melo Park California: Benjamin Cummings Publishing Company, Inc.
- Waugh R., Bonar N., Baird E., Thomas B., Graner A., Hayes P. y Powell W. 1997. Homology of AFLP products in three mapping populations of barley. *Mol. Gen. Genet.* 255: 311-321.
- Weber J.L. 1990. Informativeness of human (dC-dA) (dG-dT) polymorphism. *Genomics* 7: 524-530.
- Weeks D.E. y Lange K. 1987. Preliminary ranking procedures for multilocus ordering. *Genomics* 1: 236-242.
- Weir B.S. 1996. *Genetic Data Analysis II*. Sinauer Associates, Sunderland, MA.
- Wheeler A. 1992. A list of the common and specific names of the fishes of the British Isles. *J. Fish. Biol.* 41 (Suppl. A): 1-37.
- Williams J.G.K., Kubelik A.R., Livak K.J., Rafalski J.A. y Tingey S.V. 1990. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. *Nucliec Acid Res.* 18: 6531-6535.
- Wilson K., Li Y., Whan V., Lehnert S., Byrne K., *et al.* 2002. Genetic mapping of the black tiger shrimp *Penaeus monodon* with amplified fragment length polymorphism. *Aquaculture*, 204: 297-309.
- Windemuth C., Simianer H. y Lien S. 1998. Fitting genetic mapping functions based on sperm typing: results for three chromosomal segments in cattle. *Anim. Genet.*, 29: 425-434.
- Wintero A.K., Fredhold M. y Thomsen P.B. 1992. Variable (dG-dT)n(dC-dA)n sequences in the porcine genome. *Genomics* 12: 281-288.
- Wright J.M. 1993. DNA fingerprinting in fishes. In: Hochachka P.W., Mommsen T., (Eds.), *Biochemistry and Molecular Biology of Fishes*. Elsevier, Amsterdam, pp. 58-91.
- Young W.P., Wheeler P.A., Coryell V.H., Keim P. y Thorgaard G.H. 1998. A detailed linkage map of rainbow trout produced using doubled haploids. *Genetics* 148: 839-850.

- Yu Z. y Guo X. 2002. A basic AFLP linkage map for the eastern oyster, *Crassostrea virginica* Gmelin (abstract). *J. Shellfish Res.* 21 (1): 382.
- Yu Z. y Guo X. 2006. Identification and mapping of disease-resistance QTLs in the eastern oyster, *Crassostrea virginica* Gmelin. *Aquaculture* 254: 160-170.
- Zane L., Bargelloni L. y Patarnello T. 2002. Strategies for microsatellite isolation: a review. *Mol. Ecol.* 11: 1-16.
- Zhao H. y Speed T.P. 1996 On genetic map functions. *Genetics* 142: 1369-1377.
- Zhao L.P., Thompson E. y Prentice R. 1990. Joint estimation of recombination fractions and interference coefficients in multilocus linkage analysis. *Am. J. Hum. Genet.* 47: 255-65.

VII. APÉNDICE

Apéndice I

Extracción de ADN de alto peso molecular a partir de rodaballos diploides adultos

1. Las muestras de tejido se introdujeron en tubos estériles de 1,5 ml de capacidad junto con: 900 µl de tampón de extracción SSTNE (Blaquer, 1999), 90 µl de SDS (Sodium Dodecil Sulfato) y 4,5 µl de proteinasa K (20 mg/ml en el stock)

El tampón de extracción SSTNE incluye:

NaCl	0,30 mM
TRIS-BASE	0,05 mM
EDTA	0,20 mM
EGTA	0,20 mM
Espermidina	0,50 mM
Espermina	0,10 mM

2. Para la digestión de tejidos, los tubos se incubaron durante toda la noche a 56°C en rotación suave.
1. Al día siguiente, se añadieron 2 µl de RNAsa libre de DNAsa (10mg/ml en el stock; Roche Diagnostics) y se incubaron a 37°C durante al menos una hora.
2. Para la purificación de ADN, se añadió un volumen de fenol: cloroformo:alcohol isoamílico (24:24:1) en cada tubo. La mezcla se incubó en rotación durante 10 minutos a temperatura ambiente.
3. Concluido el tiempo de incubación, las fases acuosa y orgánica se separaron por centrifugación a 13000 r.p.m durante 15 minutos.
4. La fase acuosa se transfirió a tubos limpios y en estos tubos se repitieron los pasos 4 y 5.
5. A la fase acuosa resultante se le añadió un volumen de cloroformo: alcohol isoamílico (24:1).
6. La mezcla se centrifugó a 13000 r.p.m durante 15 minutos.
7. La fase acuosa se transfirió entonces a un tubo limpio.

8. Para la precipitación del ADN se añadió 1ml de etanol puro muy frío y ambas fases se mezclaron por inmersión. El ADN resulta entonces visible y el alcohol se puede eliminar por decantación.
9. El ADN se lavó con etanol al 70% durante una hora a 4°C. Concluida la incubación, la mezcla se centrifugó a fin de precipitar el ADN al fondo del tubo. Entonces, se eliminó el etanol por decantación. Los restos de alcohol se eliminaron por evaporación, dejando los tubos destapados a temperatura ambiente durante aproximadamente 10 minutos. El ADN libre de alcohol se resuspendió en 50 µl de TE (0,01 M Tris-HCl pH 7,4; 2,5 mM EDTA pH 8,0).

Apéndice II

Extracción de ADN de alto peso molecular a partir de embriones haploides de rodaballo

1. Cada embrión se introdujo en un tubo estéril de 1,5 ml de capacidad que contenía: 150 µl de tampón de extracción SSTNE (Blaquer, 1999), 15 µl de SDS (Sodium Dodecil Sulfato) y 5 µl de proteinasa K (20 mg/ml en el stock).

Para facilitar la digestión de los tejidos, cada embrión, una vez introducido en el tubo, fue perforado con un alfiler estéril.

2. Para la digestión enzimática, los tubos se incubaron durante toda la noche a 55°C en movimiento de rotación suave.
3. Al día siguiente se añadieron 2 µl de proteinasa K (20mg/ml en el stock) y se continuó la incubación a 55°C durante 1h.
4. La mezcla de digestión se transfirió a tubos estériles en cuyo fondo se habían depositado previamente 300 µl de Phase Lock Gel (Eppendorf).
5. A cada tubo se le añadieron 300 µl de fenol: cloroformo: alcohol isoamílico (24:24:1). Las tres fases se mezclaron por inmersión a temperatura ambiente.
6. Los tubos se centrifugaron a 13000 r.p.m durante 15 minutos
7. Durante la centrifugación, el PLG migra para formar una estrecha barrera entre las fases acuosa y orgánica. Esta separación permite un completo y fácil pipeteo de la fase acuosa

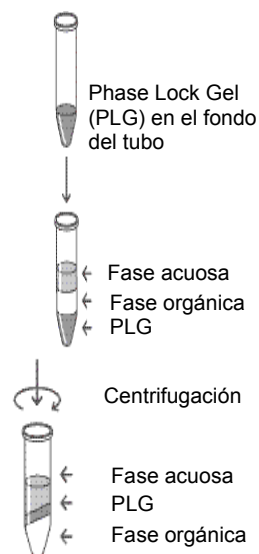


Figura VII.1. Separación de las fases acuosa y orgánica usando Phase Lock Gel (PLG). Modificado de www.eppendorf.com.

que se transfiere entonces a un nuevo microtubo estéril.

8. A la fase acuosa se le añadieron 2 µl de RNAsa libre de DNAsa (10mg/ml en el stock, Roche Diagnostics) y se incubó a 37°C durante al menos 1h.
9. Para la precipitación del ADN, se añadieron 600 µl de etanol puro frío. Los microtubos se mantuvieron a -40°C durante 3h.
10. Una vez precipitado el ADN, los tubos se centrifugaron a 13000 r.p.m durante 15 minutos a una temperatura de 4°C.
11. El alcohol se eliminó por decantación y se mantuvieron los tubos abiertos, a temperatura ambiente, hasta asegurar una completa evaporación de los restos de etanol. El ADN se hidrató con 20 µl de TE (0,01 M Tris-HCl pH 7,4; 2,5 mM EDTA pH 8,0).

Apéndice III

Protocolo de amplificación genómica GenomiPhi™

1. En un tubo de microcentrífuga (0,5 ml de volumen), se introducen 9 µl del tampón de muestra que contiene los cebadores hexaméricos para la amplificación aleatoria. En el mismo tubo se añaden 10 ng de ADN genómico extraídos mediante el protocolo de extracción fenol: cloroformo: alcohol isoamílico.
2. Se desnaturaliza el ADN a 95°C durante 3 minutos.
3. La muestra desnaturalizada se mantiene en hielo y se mezcla con 9 µl de tampón de reacción y 1 µl de mezcla enzimática. El tampón de reacción contiene las sales y los desoxinucleótidos a un pH óptimo para la reacción de síntesis de la ADN polimerasa *Phi29*.
4. La mezcla de reacción se incuba a 30°C durante toda la noche (16-18h)
5. Finalmente, se somete la mezcla a 65°C durante 10 minutos a fin de inactivar a la polimerasa *Phi29*.

Apéndice IV

Protocolo de amplificación PCR de los loci microsatélite

Mezcla de reacción: Cada PCR se llevó a cabo en un volumen total de 15 µl conteniendo:

- 30 ng de ADN genómico
- 1 µM de cada cebador
- 0,1 mM de dATP, dTTP, dCTP, dGTP
- tampón 1,5 x PCR gold buffer (15 mM Tris-HCL, pH 8,0; 50 mM KCl)
- 1,5 MgCl₂
- 0,5 unidades de ADN polimerasa *AmpliTaq Gold*.

Programa de la PCR de gradiente:

- desnaturalización a 95°C, 10 minutos.
- 35 ciclos:
 - 94°C, 45 segundos
 - T_a, 50 segundos
 - 72°C, 50 segundos
 - 72°C, 10 minutos

Apéndice V

Protocolo AFLP

Restricción-Ligación

Aproximadamente, 100 ng de ADN genómico se digirieron inicialmente durante 1h 30 min a 65°C en un volumen de reacción de 10 µl que incluía: 1 U de *TaqI* endonucleasa; tampón 1x (“one phor al buffer plus”, Amersham Pharmacia Biotech); 10 ng/ µl de BSA (Bovine Serum Albumin); y 1 mM de DTT (Ditiotreitol).

Transcurrido el tiempo de digestión con *TaqI*, a los 10 µl iniciales se añadieron 6 µl de una mezcla compuesta por: 1U de *EcoRI* endonucleasa; tampón 1x (“one phor al buffer plus”, Amersham Pharmacia Biotech); 10 ng/ µl de BSA (Bovine Serum Albumin); y 1 mM de DTT (Ditiotreitol). En este volumen final de 16 µl se produjo la digestión con *EcoRI* a 37°C durante 2h.

El ligamiento de los adaptadores sintéticos a los fragmentos de restricción se llevó a cabo durante 16-18h a 16 °C añadiendo a los 16 µl de solución de restricción 4 µl de una solución que contenía: 0,2U de T4 DNA ligasa; adaptadores *EcoRI* (1 pmol) y adaptadores *TaqI* (10 pmol); 0,2 mM de ATP (adenosin trifosfato); tampón 1x “one phor al buffer plus”; 10 ng/ µl de BSA; y 1mM de DTT (Ditiotreitol).

Amplificación AFLPs

La amplificación de los fragmentos de restricción se llevó a cabo en dos pasos: una primera preamplificación con cebadores que incluían una base selectiva y una segunda amplificación del preamplificado mediante cebadores con 3 bases selectivas. La mayoría de los protocolos, incluyendo el de Papa *et al.* (2005) incorporan estas amplificaciones en dos pasos para minimizar artefactos en el perfil AFLP de genomas complejos.

i) *Preamplificación.* Para la preamplificación se utilizaron cebadores complementarios a la secuencia del adaptador incluyendo un nucleótido

selectivo. El cebador *EcoRI* que lleva la adenina (A) como base selectiva, se combinó con un cebador *TaqI* llevando adenina (A) o citosina (C) como bases selectivas. Diez µl de ADN ligado (aproximadamente 50 ng de ADN molde) se amplificaron en 20 µl de un volumen de reacción que contenía: cantidades iguales de cebadores *EcoRI* y *TaqI* (30 ng); 0,2 mM de cada uno de los cuatro dNTPs; 1,5 mM MgCl₂; 10x tampón (10 mM Tris-HCl pH 8,3; 50 mM KCl); y 0,4 U *Taq* polimerasa (Roche Molecular Biochemical, Mannheim Germany). El ADN molde se amplificó según el protocolo PCR de Papa *et al.* (2005):

- 94°C, 2 minutos
- 30 ciclos:
 - 94°C, 30 segundos
 - 56°C, 1 minuto
 - 72°C, 1 minuto
- 72°C, 7 minutos.

Este paso de preamplificación proporciona una cantidad prácticamente ilimitada de ADN molde para las posteriores amplificaciones con pares de cebadores de tres bases selectivas.

ii) *Amplificación*. El ADN preamplificado a partir de cebadores *EcoRI*-A y *TaqI*-A se usó como molde para las amplificaciones con cebadores *EcoRI*-ANN/*TaqI*-ANN (N representa cualquiera de las 4 bases posibles) y el ADN preamplificado a partir de los cebadores *EcoRI*-A y *Taq*-C se usó como molde para las amplificaciones con las combinaciones *EcoRI*-ANN/*TaqI*-CNN.

Los cebadores *EcoRI*-ANN se marcaron con fluorescencia (Cy5) para, posteriormente, permitir la detección automatizada de los fragmentos amplificados. Del ADN preamplificado se usaron 0,3 µl como molde para la amplificación selectiva en 20 µl de un volumen de reacción que contenía: 10 ng de cebador *EcoRI* marcado (Cy5); 30 ng de cebador *TaqI* marcado; 0,2 mM de cada uno de los cuatro dNTPs; 1,5 Mm MgCl₂; 10x tampón (Roche Molecular Biochemical, Mannheim Germany); y 1U de *Taq* polimerasa (Roche Molecular

Biochemical, Mannheim Germany). La amplificación PCR transcurrió siguiendo el protocolo de Papa *et al.* (2005):

- 94°C, 2 minutos
- 13 ciclos de amplificación “touch down”:
 - 94°C, 30 segundos
 - 65-56°C, 30 segundos (disminuyendo 0,7°C en cada ciclo)
 - 72°C, 2 minuto
- 30 ciclos:
 - 94°C, 30 segundos
 - 56°C, 30 segundos
 - 72°C, 2 minuto
 - 72 °C, 7 minutos.

Todas las reacciones de PCR se llevaron a cabo en una PT-200 (MJ Research).

VIII. *ABREVIATURAS*

AFLP	Amplified Fragment Length Polymorphism
AT	Adenina, Timida
C	Coeficiente de coincidencia
Ce	Cobertura genómica esperada
Co	Cobertura genómica observada
CEPH	Centre d'Etude du Polymorphisme Humain
chi²	Test chi cuadrado
cM	Centimorgan
ESTs	Expressed Sequence Tags
F₁	Primera generación filial
F₂	Segunda generación filial
FIASCO	Fast Isolation by AFLP of Sequences Containing repeats
GC	Guanina, Citosina
Ge	Longitud genómica esperada
Go	Longitud genómica observada
He	Heterozigosis esperada
Ho	Heterozigosis observada
HWE	Equilibrio Hardy-Weinberg
<i>i</i>	Interferencia
LG	Grupo de ligamiento
P	Generación parental
PCR	Polymerase Chain Reaction
PIC	Polymorphic Information Index
QTL	Quantitative Trait Loci
<i>r</i>	Frecuencia de recombinación
R	Coeficiente de correlación de Pearson
RIL	Recombinant inbred lines
RFLP	Random Fragment Length Polymorphism
RAPD	Random Amplified Polymorphic DNA
SNP	Single Nucleotide Polymorphism

SSR	Single Sequence Repeat
STR	Single tandem repeat
T_a	Temperatura de anillamiento